Influential Spreaders in the Political Twitter Sphere of the 2013 Malaysian General Election

Hong-liang Sun * + [1], Eugene Ch'ng * + [2], and Simon See [3]

International Doctoral Innovation Centre, the University of Nottingham Ningbo China, Science and Engineering Building, 199 Taikang East Road, 315100 Ningbo, China *
zx17898@nottingham.edu.cn [1]
NVIDIA Joint-Lab on Mixed Reality, the University of Nottingham Ningbo China+
eugene.chng@nottingham.edu.cn [2]
NVIDIA Technology Centre APJ ssee@nvidia.com [3]

Abstract

Purpose – The article investigates political influential spreaders in Twitter at the juncture before and after the Malaysian General Election in 2013 (MGE2013) for the purpose of understanding if the political sphere within Twitter reflects the intentions, popularity and influence of political figures in the year in which Malaysia has its first 'social media election'.

Design/methodology/approach – A Big Data approach was used for acquiring a series of longitudinal data sets during the election period. The work differs from existing methods focusing on the general statistics of the number of followers, supporters, sentiment analysis and etc. A retweeting network has been extracted from tweets and retweets and has been mapped to a novel information flow and propagation network we developed. We conducted quantitative studies using k-shell decomposition, which enables the construction of a quantitative Twitter political propagation sphere where members posited at the core areas are more influential than those in the outer circles and periphery.

Findings – We conducted a comparative study of the influential members of Twitter political propagation sphere on the election day and the day after. We found that representatives of political parties which are located at the center of the propagation network are winners of the presidential election. This may indicate that influential power within Twitter is positively related to the final election results, at least in MGE2013. Furthermore, a number of non-politicians located at the center of the propagation network also significantly influenced the election.

Research limitations/implications – This research is based on a large electoral campaign in a specific election period, and within a predefined nation. While the result is significant and meaningful, more case studies are needed for generalised application for identifying potential winning candidates in future social-media fueled political elections.

Practical implications – We presented a simple yet effective model for identifying influential spreaders in the Twitter political sphere. The application of our approach yielded the conclusion that online 'coreness' score has significant influence to the final offline electoral results. This presents great opportunities for applying our novel methodology in the upcoming Malaysian General Election in 2018. The discovery presented here can be used for understanding how different players of political parties engage themselves in the election game in Twitter. Our approach can also be adopted as a factor of influence for offline electoral activities. The conception of a quantitative approach in electoral results greatly influenced by social media means that comparative studies could be made in future elections.

Originality/value – Existing works related to general elections of various nations have either bypassed or ignored the subtle links between online and offline influential propagations. The modeling of influence from social media using a longitudinal and multilayered approach is also rarely studied. This simple yet effective method provides a new perspective of practice for understanding how different players behave and mutually shape each other over time in the election game.

1. Introduction

The advent of social media has provided individuals traditionally confined to local interactions a means to reach out globally. Autonomous and independent decisions for voicing out opinions, to publish and to subscribe to channels are some of the freedoms that social media provides. Twitter is one such public platform where even 140-more-or-less characters wide service can link people and spread information across the world. This simple micro-blogging platform has contributed huge impacts to billions of individuals in various needs and events.

The phenomenal data generated within Twitter has also drawn attention to extensive studies within the field of *computational social science*. Surveys conducted by social scientists were limited prior to the availability of social media such as those accomplished by a small number of college students (Giles, 2012). With the advent of the emerging social media landscape, a significant change in how data is processed and analyzed can be witnessed, particularly since the new era of social media bringing the social sciences into a data rich and quantified environment. Due to the accessibility and the ubiquity of social media data (Ngai, Moon, Lam, Chin, & Tao, 2015), they have been used for forecasting forthcoming events such as consumer product demands (Chong, Ch'ng, Liu, & Li, 2015), online product sales (Chong, Li, Ngai, Eugene, & Lee, 2015), social networking sites continuance intention (Christopher, Jeff, & Judy, 2016), box-office revenue of movies (Sitaram & Huberman, 2010), video games sales (Goel, J., & Lahaie, 2010), stock market fluctuations (Johan, Mao, & Zeng, 2010) and etc.

However, Twitter has not been fully explored, especially for predicting electoral campaigns with quantitative approaches. Most studies are exploratory and there are no general agreements from numerous researchers on measures leading to the prediction of political election (Metaxas, Takis, & Mustafaraj, 2012). Different strategies including user behaviour, Twitter specific features (retweet or reply), sentiment analysis from tweets content can be applied to infer political trends (Dwi, Nugroho, & Hauff, 2015). A better approach by Wong et al. studied the problem of quantifying and inferring the political leaning of Twitter users in 2012 US Presidential Election. Their method achieved good correlations as compared to manually created labels. Wong et al.'s quantitative study provided insightful diagrams on political demographics of the Twitter population (Wong, Tan, Sen, & al., 2016). Boutet et al. analysed the characteristics of the three main parties in 2010 UK General Election. They proposed a classification algorithm to identify the political leaning of users using the amount of tweets and retweets closely related to political parties. The results achieved 86% classification accuracy referring to a particular political party (Boutet, Kim, & Yoneki, 2012). Tumasjan et al. studied the tweets' political sentiment during the German Federal Election in 2009. An analysis of the tweets' political sentiment demonstrated close correspondence to the politicians' political positions, which indicated that the online content of Twitter messages can reflect the offline political landscape (Tumasjan, Sprenger, & Sandner, 2010). Bermingham et al. combined sentiment analysis using supervised learning and volume-based measures to capture the voting intentions in Irish General Election in 2011. They found that social analytics using volume-based measures and sentiment analysis were predictive and they made a number of observations related to the task of monitoring public sentiment (Bermingham & Smeaton, 2011). Although exploratory, these are notable studies. These studies are mainly focused on Twitter user behaviours and the sentiments of tweets, studies on the correlations between politician online influence in terms of network centrality and offline electoral results are scarce. In this paper we aim to study the interplay between online influential spreaders and final electoral results in 2013 Malaysia's general election using network coreness centrality.

Opportunities for a deeper and more interactive approach to media presented itself in Twitter, which brings us closer to Jurgen Habermas's idea of the public sphere, a sphere that permits citizens to interact, debate on the public issues without fear of political power (Beers, 2006). In the 12th Malaysia's general election in 2008, Twitter opened up a new platform for Malaysian citizens to deliberate discourse on political issues, during which the government underestimated the influence of social media from voters (Macnamara & Ansgar, 2013). The National Front (Barisan National or BN), which was fully supported by mainstream media of Malaysia, suffered losses to opposition's coalition The Parti Rakyat Malaysia (PR) as a result. PR is a loose coalition composing of Democratic Action Party (DAP), People Justice Party (PKR – Pakatan Rakyat) and the Pan-Malaysian Islamic Party (PAS). To acknowledge the power of social media, the Prime Minister of Malaysia Tun Abdullah Ahmad Badawi, admitted on 25th March 2008 that BN had lost the online war with the oppositions in 2008 due to underestimating the power of online social media. The year 2008 appears to be the year where online contents have become catalysts to political influence, leading towards the 2013 with Malaysia's first 'social media election' (Lim, 2013). Furthermore, he pointed out that BN gave more priority to traditional media than social media, which led to painful losses (i.e. lost the states of Penang, Selangor, Kedah, Perak, and Kelantan) (Lalitha, 2013). From the perspective of the opposition, having limited exposure from traditional mass media, PR depended solely on social media to disseminate information or to organize offline activities, which contributed to greater support from participants online.

We are curious to know if social media can indeed impact the electoral outcome in Malaysia's 13th general election in 2013 (MGE2013). Hence, our aim is to study how different politicians and citizens conducted themselves in various personas in the Twitter political sphere during the MGE2013. Can online activities in terms of tweets genuinely reflect offline behaviors of different players in the election game? Previous studies by Goncalves et al. have demonstrated that the volume of tweets in Twitter is strongly correlated with offline events in the real world (J., Goncalves, Ramasco, & Cattuto, 2012). Furthermore, Twitter usage in the 2011 Spanish local mayor election demonstrates that it is used as a single-directional tool to disseminate news and information while its bi-directional discussions are ignored (Criado, Ignacio, Martínez-Fuentes, & Silván, 2012). With these as a backdrop, we began our focus on Twitter data from the 2013 Malaysian General Election. This has, as a basis from which we scope our work, drawing upon the conclusion from the Prime Minister of Malaysia that MGE2013 can be considered as Malaysia's first "social media election" (Lim, 2013).

In this work, we have particular interests on how well Twitter activities can reflect the intentions, popularity and influence of political figures in Malaysia. Our initial thought is that the tweets of leading politicians will generate many retweets. We also believe that the retweets of followers of the politicians who themselves are not politicians will be highly retweeted as well. Are the tweets of perceived influential political leaders retweeted more as compared to tweets made by the non-politicians? The answer to this question will reflect how different individuals (politicians, non-political public figures) engage in the Twitter sphere of MGE2013. On the other hand, more attention will need to be focused on understanding the predictive power in terms of Twitter network centrality. Will online social media influence with respect to network centrality positively correlate with offline election results? Leading politicians from different political parties get their tweets retweeted broadly to disseminate information related to MGE2013. Extensive propagation of such messages as the form of tweets and retweets may trigger deep and broad discussions among voters was our hypothesis. It is for this reason that we have sought to investigate the interplay between online influence in the form of network centrality and offline election results.

The remainder of this paper is presented in the following order. We formulate research questions related to Malaysian political Twitter sphere in the election in the next section, drawing from our discussions in the previous paragraph. We then continue with a description of how we collected data and generated a series of longitudinal retweeting networks, these identified the most influential spreaders in Twitter during the election campaigns. We believe that our method will reveal the relationships between the most influential

spreaders and political parties and voters in the "first social media election in Malaysia".

2. Methodology

The provision of 140-more-or-less tweets can support the open political communication environment during election campaigns. Here, we base our hypotheses and introduce the methods on the specific political settings in the 2013 Malaysian General Election.

2.1 Data Descriptions

The big data software that we have used to collect and collate our Twitter data is installed on a total of 16 Linux 64-bit Ubuntu (12.04) Virtual Machines (VM) within four (4x) Dell PowerEdge C6100. It was found to be sufficient for the data captured, for storage in the distributed database Mongo DB servers and processed in real-time, as data streams using Node.js applications we developed. Data have been collected using the Twitter streaming API with keywords as topical filters. Captured data stream is distributed over three Mongo DB shards with JSON key-value pairs (Ch'ng, 2014).

This study used data from tweets collected from April 22 2013 to April 28 2013, and from May 1 2013 to May 7 2013. The data set consists of 1,122,791 tweets from Malaysian Twitter users living in Malaysia or abroad. Besides general users identified with their own Twitter accounts, well-known politicians or official accounts of political parties have been identified and are listed in Table 1. All tweets were collected using a group of predefined hash tags including "#GE13, #PRU13" et al. Most influential leaders of main political parties are involved. Barisan Nasional (BN) has been the ruling party in Malaysia for a long period, since the year 1974 to be specific. The main opposition parties consist of People's Justice Party (PKR – Pakatan Rakyat), Democratic Action Party (DAP) and Pan-Malaysian Islamic Party (PAS).

Table 1 Political parties and their representative Twitter accounts in MGE2013

2.2 Method

Different from the conventional method of mapping follower-followee networks, our approach defined here maps instantaneous activities over a series of days closely related to MGE2013. This is more meaningful as tweets define the true interactions of active users on Twitter, not just followers (Ch'ng, 2015),

As mentioned earlier in data descriptions, tweets were collected from April 20 2013 to April 28 2013, and from May 1 2013 to May 7 2013. This has actually resulted in 69 longitudinal datasets, each of which contains 5 hours of continuous datasets. We have particular interests in the comparison study of information diffusion before and after the election day.

In our method, the clustering of opinions is the first step in identifying popular accounts. Our method divides the propagation network into different communities where users with similar opinions form the same group and users with different opinions are segregated (Sun, Ch'ng, Yong, Garibaldi, & Chen, 2018).

Thus the propagation network is divided into communication communities, termed 'activity networks' (Ch'ng, 2015) using community detection algorithm by Blondel et al. (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). Our argument is that community discovery can enhance the understanding mutual relationships of different political parties and their closely related members.

We aim to investigate social media online influential power using influential spreaders. The identification of the most efficient spreaders in a network is an important step towards understanding the spread of opinions by voters. The influence of a node is closely related to the topological structure of the underlying network and where it is located. The use of degree centrality is a straightforward approach for counting the neighbours of a certain node. It has limitations because only directly linked neighbours are considered while neighbours of neighbours are ignored. Betweeness centrality is another centrality used for measuring the importance of a node. It is equal to the number of shortest paths from all nodes linked to all others that passed through that node. Kitsak et al. argued that the influence of a node is determined by its position of global network structure rather than its local neighbours (Kitsak, Gallos, Havlin, & Liljeros, 2010). Coreness was proposed to iteratively decompose the networks where nodes located at the central shells were more influential than those in periphery shells. Experimental studies have shown that coreness outperforms degree centrality and betweeness centrality with lower error rates in propagation simulation on real networks. For this reason, we have used coreness instead of other centrality measures to denote the importance of a node.

In contrast to common belief, the best spreaders are not the most highly connected people. It is well understood that the most influential spreaders are those locating at the core of the network using a k-shell decomposition method (Kitsak, Gallos, Havlin, & Liljeros, 2010). The k-shell decomposition method decomposes a network into hierarchically ordered shells by recursively pruning the nodes with degree less than the current shell index. It starts by removing all nodes with degree $k=1$. After which all nodes with $k=1$ should be iteratively removed until no nodes with $k=1$ are left. These initial set of nodes are assigned to the first shell with coreness $k_s=1$. Similarly, nodes with degree $k=2$ are iteratively removed and assigned to the second shell with $k_s=2$. This process continues until no nodes are left belonging to a high level shell. A dummy sample network with 10 nodes using k-shell decomposition method is shown in Figure 1(a) as an example. In the case of networks with tens of thousands nodes and edges, only selected nodes are displayed as needed. A simplified visual representation of the selected nodes is shown in Figure 1(b). The k-shell decomposition program was implemented using NetworkX in Python (https://networkx.github.io/).

Figure 1 A dummy sample of k-shell decomposition applied to a network with 10 nodes. (a) Node 1, 2, 3, 4 locate at the dark red core zone with $k_s=3$, node 5 and 6 have $k_s=2$ and node 7, 8, 9, 10 have $k_s=1$. (b) The network topological structure is ignored and nodes are assigned to different layers according to their k_s scores.

For selected Twitter users and propagation networks, we obtained the corness scores of users by using k-shell decomposition method mentioned above. Such coreness scores naturally contribute to the layer division of their propagation sphere. Thus different users are assigned to different propagation layers according to their corness scores. Selected users and their corness scores are visualised within concentric circles as respective propagation spheres. A comparative analysis of selected users and their positions within the propagation spheres contributes to answering the research questions proposed previously.

3. Results

In this section, we test our hypotheses on the Malaysian political Twitter sphere using tweets collected previously.

3.1 Retweeting Propagation Networks of MGE2013

The first step is to construct propagation networks during the election period. Only tweets containing the keyword "*#GE13 #PRU13*" were captured via the streaming API from Twitter's service and recorded in our distributed database mentioned previously.

Figure 2 shows that before May 4, the fluctuations of tweets are small and the volume is at the lowest levels. The levelness of the graph of tweets in this period indicates little events. The trends appear to change dramatically on May 4 2013, a day before the general election. The activity reaches the first peak shown in the red dot. When night falls, the number of tweets is at a low level for 5 hours. It subsequently peaks during the electoral campaign on May 5 2013. After the election result is officially announced, the number of tweets decreases dramatically to nearly the same volume prior to MGE2013. The two peaks respectively relate to the day before and the day after the MGE2013. From the fluctuations of tweets every five hours above, it can be demonstrated that the pairs of valleys and peaks refer to the day before and the election day. The highest peak indicates the time prior to the official release of the results. Thus a high peak and long tail pattern demonstrates that the online social activities in terms of tweets have heightened activity immediately after a long period of silence.

Figure 2 Number of Tweets every five hours during MGE2013

The graph shown in Figure 2 gives us interest in exploring the tweeting and retweeting activities on the election day and the day after. We believe that a comparative study of the two datasets will uncover the collective patterns of politicians, non-political public figures and voters. Tweets with high popularity will be disseminated to large groups of users via retweeting. In this paper, we focus on information diffusion patterns in terms of retweet.

A sample of retweet is listed below:

"*tweet_98788, RT @chris_chong_n Its a very smart move by the govt to abolish govt school fees! first time in Malaysia! Bravo to BN govt! #GE13 #PRU13*"
"*tweet_98788*" -- "@mansempoi82"
Here, "*tweet_98788*" is the unique identifier of the tweet added by our preprocessing program. "*RT @chris_chong_n*" indicates it is a retweet from Twitter user @*chris_chong_n.* Hash tags *#GE13 #PRU13* indicate that this is a retweet closely related to MGE2013.

From the tweet "*tweet_98788*" -- "@mansempoi82" it is also known that this is a retweet by "@mansempoi82". Thus a retweet here is represented by an edge from "@mansempoi82" to "@*chris_chong_n*" in the propagation network. We extracted two network structures at the election day of MGE2013 and one day after respectively. The basic topological features of the two propagation networks are illustrated in Table 2. We can see that tweeting and retweeting activities decreases exponentially following the election day.

Table 2 Topological features of propagation networks on the election day and the day after in MGE2013.

3.2 Information Diffusion on the Election Day

We conducted an analysis of the tweets propagation network on the election day, which provides valuable clues on how different political parties interact with their voters in the Twitter sphere. Further explanations will answer research questions related to which accounts are widely known and retweeted by voters.

Our method divides the propagation network into different communities (Jie Cao, 2016)where users with similar opinions form the same group and users with different opinions are segregated (Jie Cao Z. W., 2013). Thus a propagation network is divided into communication communities using community detection algorithms (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008).

Table 3 shows that 9745 BN fans were grouped into the same community. This community contains the Twitter account of Najib "@NajibRazak" and BN official account "@barisanasional". But its size is not as large as the opposition parties. Using our method, voters who paid close attention to PKR are separated into a large group of 19817 users and another small group consisting of 2234 users. Both DAP and PAS have their own communities, but the former is larger than the BN group and the latter is comparatively smaller. Aside from the official accounts of political parties, mass media sites also have their own communities, which are groups discussing the latest news and opinions. Here, we list three representative groups.

Table 3 Community divisions of Twitter propagation network on election day in MGE2013. C indicates the clustering coefficients, l denotes the average shortest path length, and r represents the assortativity.

Although different political parties have their own communities to disseminate particular opinions and instantaneous news within closed groups, such groups are not isolated from each other. Conversely, distinct opinions and voices from every group are merged into the whole Twitter political sphere. A better way to understand how tweets are propagated within this information sphere comes from the idea of dividing the propagation network into concentric circles. The core zone contains the most influential spreaders whose tweets are widely retweeted by users coming from periphery zones. The idea is simple but the implications for identifying influence spreaders is great.

Propagation of tweets can be modeled as a network representing how ideas spread. Identification of the most efficient spreaders in a network is valuable to understand the spread of opinions by politicians and voters. K-shell decomposition method (Kitsak, Gallos, Havlin, & Liljeros, 2010) is employed to obtain the coreness scores of selected spreaders. The higher value of corness score indicates the influential power of spreaders. Given selected spreaders and their coreness scores by k-shell decomposition, they are naturally divided into different groups. It is shown in Figure 3 and Table 4 that the most influential spreaders with coreness scocre k_s=28 are visualised at the center of propagation sphere as red core layer. Orange transition layer contains spreaders with $21<k_s<27$, light purple middle layer contains spreaders with $11<k_s<20$ and sky blue periphery layer consists spreaders with $1<k_s<10$.

It is discovered in Figure 3 and Table 4 that BN is located in the middle of the red core zone while loosely connected alliance of PKR, DAP and PAS have different positions in the orange transition layer, the light purple middle layer and the sky blue periphery layer. Social influence can be described as a form of ability

where there is sustained influence of a person on the thoughts or actions of others. In social networks such as Twitter, if a message posted by Bob is retweeted by Alice, then Alice is, in a sense influenced by Bob in terms the need to retweet. K-shell decomposition is a method to measure the influence of users in online social networks such as Twitter and etc. Examples can be seen in Brown et al., who identified a small number of most important users from a large-scale Twitter network with roughly 7 million users and more than 80 million actual tweets (Brown, 2011), and Kwak et al. who identified the top 20 influential users from a Twitter social network with 41.7 million users and 1.47 billion relationships (Kwak, 2010). Here we sought to rank the online influence of political figures from BN and its opposing parties using k-shell decomposition. Experimental results showed that political figures in BN are ranked higher than politicians from its opposing parties using coreness scores before and after Malaysia's General Election in 2013. This implies that politicians from BN are more influential than those from the opposition. This results in terms of online influence does fit the final election results quite well.

Another discovery from this political tweet sphere is that a number of non-politician Twitter accounts appear in the core zone. It is for the purpose of individual privacy protection that the middle characters of the names are replaced by "***". For instance, "@K***N" is a Twitter account of an independent user whose novel voice gives rise to extensive propagation. Such independent voices occupy more positions as compared to politician accounts in the center of propagation sphere, which leads to answer the research questions formulated in the previous section. This is shown in Table 5 where there were six influential spreaders supported BN, four supported PR and three others whom paid more attention to pubic affairs. There were clear symbols from the tweets to support BN (i.e. #Vote4BN) or to support PR (God willing PR can win). It is also interesting to see public affairs such as housing issues and school fees as heightened concerns amongst the influential spreaders. Such influential spreaders display Jurgen Habermas's ideal of the public sphere, especially in the General Election of Malaysia, a sphere that permits citizens to interact and debate on public issues without fear of political powers (Beers, 2006). We observed that in MGE2013, Twitter is not a tool that are merely employed to disseminate information to voters by politicians as in the case of Spanish local mayor elections (Criado, Ignacio, Martínez-Fuentes, & Silván, 2012). However, a large number of independent citizens do engage and play significant roles as influential opinion spreaders. Effective use of Twitter by politicians for the purpose of discussing well-concerned social issues with voters can potentially have influential effects on the final political directions of the nation.

Figure 3 The Twitter propagation sphere on the election day of MGE2013. Personal Twitter account names have been made anonymous to protect privacy, but public figures remain public.

Table 4 Coreness scores of selected Twitter accounts on the election day of MGE2013. Core (coreness equals 28), transition (coreness ranging from 21 to 27), middle (coreness ranging from 11 to 20) and periphery (coreness ranging from 1 to 10).

Table 5 Influential Spreaders and their tweets.

3.3 Information Diffusion after the Election Day of MGE2013

Despite facing unprecedented pressure from the opposing alliance of PKR, DAP and PAS, BN was returned to the center of political power stage on May 5 2013 after the uncertain electoral campaign. It is shown in Table 5 that the major members of BN are maintaining a cohesive social community after the election day even though the scale of the community have shrunk in terms of node and edge numbers compared with the data in Table 3. It is common sense to note that some of the voters eventually leave the political communication circle once the results of MGE2013 are made public. Differing from BN, the members of PKR are merged into a small community, the members of DAP are split into two small communities and PAS actually joins the community of BN. Due to the new power entitled to BN, its opposing alliance was split into different groups. However, they appeared to be involved in the political sphere – they were restating their political directions to their supporters even when the political game was approaching to the end.

Table 6 Community divisions of Twitter propagation network on the election day in MGE2013. C indicates the clustering coefficients, l denotes the average shortest path length, and r represents the assortativity.

Further inspections on selected Twitter accounts of different political parties uncover changes in the propagation sphere. The community partitions of different political figures are shown in Table 6. It is shown in Table 7 that the largest coreness score, which has decreased from 28 in Table 4 to 26 here. This indicates that the total size of the sphere has shrunk after the official declaration of the final results of MGE2013. BN retained the core of the new administration, and is still located at the center of the Twitter political sphere (Figure 4). However, the representative accounts of the opposition alliance which include "@anwaribrahim", "@limkitsiang", "@ tuan_ibrahim" and etc. demonstrated that they have swiftly lost the focus and have been swiped out of the middle and the transitional layers. The coreness scores becoming less than 10 demonstrate that their messages are now no longer widely propagated. In spite of the propagation priorities of politicians in Malaysia, a further examination shows that the focus now switches to non-politicians. From comparisons made of accounts located at the core layer (between Figure 4 and Figure 3), we can see that exchanges between core layers and transition layers occur with respect to independent yet pioneer non-politicians. "@z***i" and others have moved out of the core layer into the transition layer, but fresh non-politicians such as "@L***2" and others have now entered the core area of the political sphere. Some of the prior non-politicians (@D***i and others remained stable across the entire election period. These changes of focus to non-politicians within the core layers indicated a public swap of interest after the electoral results.

Across the entire election period of MGE2013, it is of particular interests to us that the pioneering non-politicians have entered the center of the Twitter political landscape. Although politicians think the dissemination of as many messages as possible so as to persuade voters to support their ideas in national affairs, it is the non-politicians' voices that have been extensively broadcasted to the mass media. Micro-blogging platform Twitter can certainly provide an open, democratic and effective way for expressing public concerns for both politicians and non-politicians.

Table 7 Coreness scores of selected Twitter accounts after the election day of MGE2013. Core (coreness equals 26), transition (coreness ranging rom 21 to 25), middle (coreness ranging rom 11 to 20) and periphery (coreness ranging rom 1 to 10).

Figure 4 Tweets propagation sphere after the election day of MGE2013

4. Discussions

In this paper, a Big Data Twitter architecture developed by our group (Ch'ng, 2014) was used for recording a longitudinal dataset associated with the Twitter hashtags *#GE13 and #PRU13* during the Malaysian General Election 2013. Our activity-based method differs from existing methods, which analyse the number of followers on the main profile pages of specific Twitter accounts, number of tweets, sentiments trends, and etc. In our method, retweeting networks was extracted from tweets, followed by the application of community detection algorithms to investigate political communication patterns within the tweets. We further applied a novel k-shell decomposition method to construct a political propagation Twitter sphere consisting of different layers. The new methods presented here have uncovered the propagation rules of politicians from the main political parties – BN, PKR, DAP, PAS, and non-politicians (active citizens) on the propagation sphere of MGE2013. Further analysis and synthesis in the political propagation sphere on Twitter have provided clues for answering the research questions proposed in our study.

The first research question investigates whether tweets by politicians are retweeted more as compared to non-politicians. The answer to this question determines whether Twitter is used in MGE2013 merely as a tool for information dissemination of political needs. Given the retweeting networks captured before and after the election, we constructed two Twitter propagation spheres showing that more influential spreaders are located in the inner cores and less influential spreaders are located at the periphery using a k-shell decomposition method. From the propagation sphere we can see that most members located at the central areas are non-politicians (opinion leaders) and BN official accounts are always in the core areas. We conclude that non-politicians obtained more influential power than politicians in Twitter in the MGE2013. In this particular election, Twitter is not a tool that is solely employed to disseminate information to voters by politicians as in the case of Spanish local mayor elections (Criado, Ignacio, Martínez-Fuentes, & Silván, 2012). However, a large number of independent citizens engaged and played significant roles as influential opinion spreaders in MGE2013.

The second research question is related to the prediction power of online tweets and offline votes. Is the Twitter influential power of political parties positively related to final election results? We compared the positions of various political parties before and after the MGE2013 to see if influential power on social media indeed correlates with the final results of MGE2013. This has been verified due to the gaps between BN and its opposing alliance of PKR, DAP and PAS. After the official announcement of the election results, BN still appears to be located at the core areas of Twitter political sphere while loosely connected opposing alliance transitions into the periphery. We conclude that online influential power of different political parties can positively relate to the final election result. Our method for identifying influential spreaders is a good predictor of the MGE2013.

In summary, a political product of two factors, politicians and independent citizens, have been investigated to understand how open micro-blogging services such as Twitter can contribute to a democratic public affairs of elections, particularly the social media-fueled Malaysian General Election in 2013. Apart from prominent politicians, we have observed that a number of non-politicians occupied the core areas of the propagation landscape. This specific phenomenon in Malaysia demonstrates a true reflection of how independent citizens truly and actively engaged with the general election in the 2013. Their voices have been heard by the masses because of their central positions in the propagation landscape

both before and after the election day. Twitter can indeed provide an open, democratic and effective platform from which both Malaysian citizens as well as politicians can create influence, especially through the active communication of their opinions and ideas. Our studies have also shown that politicians appeared to not have as much impact as expected, as compared with when they used and influenced traditional media. This discovery indirectly supports the proposal that the MGE2013 is a "first social media election in Malaysia" (Lim, 2013).

The novel method we have investigated displays the political ecosystem of the various layers of Malaysia's general election in 2013. We strongly believe that social media will continue to be crucial for the ruling party, opposition coalitions and voters in the interim years leading to the next elections. However, real political issues in essence will still be the major determinant for final election results, in combination with social media. Our research provides a new way to display Jurgen Habermas's ideal of the public sphere, especially in the General Election of Malaysia, a sphere that permits citizens to interact, and debate on the public issues without fear of political powers (Beers, 2006). It is also meaningful to model similar public issues of other nations in the future.

Our study is with limitations as it has only been tested in a specific general election and within one social media platform. As it is, Twitter, whilst highly important in the MGE2018 is not the only social media where the ruling party BN and the opposition's coalition PR attempts to compete and organize their online campaigns. Other social media platforms such as Facebook, YouTube, Instagram and Snapchat can also play crucial roles in information propagation during the electoral period. A better model capturing data from multiple media sites are being looked into for the up-coming General Election of Malaysia in 2018.

Acknowledgements

References

Albert, F., M., K., Sprenger, T. O., & Welpe, I. M. (2011). Divided They Tweet: The Network Structure of Political Microbloggers and Discussion Topics. *Proceedings of the Fifth International AAAI Conferences on Weblogs and Social Media (ICWSM 2011)*(Barcelona, July 17-21, 2011)).

Alex, B., & Eltham, B. (2009, November 19-20). Twitter free Iran: An evaluation of Twitter's role in public diplomacy and information operations in Iran's 2009 election crisis. *Record of the Communications Policy & Research Forum 2009* , 322-334.

Beers, D. (2006). The Public Sphere and Online, Independent Journalism. *Canadian Journal of Education*, 29 (1), 109-130.

Bermingham, A., & Smeaton, A. (2011, 2 10). On using Twitter to monitor political sentiment and predict election results. *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011)*. , pp. 1-10.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008 (10), 10008.

Boutet, A., Kim, H., & Yoneki, E. (2012, 12). What's in Your Tweets? I Know Who You Supported in the UK 2010 General Election. *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, pp. 411-414.

Brown, Phil E., and Junlan Feng. Measuring user influence on twitter using modified k-shell decomposition. Fifth international AAAI conference on weblogs and social media (Barcelona, Spain) 2011.

Ch'ng, E. (2015). Social Information Landscapes: Automated Mapping of Large Multimodal, Longitudinal Social Networks. *Industrial Management & Data Systems*, 115 (9), 1724-1751.

Ch'ng, E. (2015). The bottom-up formation and maintenance of a Twitter community: Analysis of the #FreeJahar Twitter community. *Industrial Management and Data Systems*, 115 (4), 612-624.

Ch'ng, E. (2014). The Value of Using Big Data Technology in Computational Social Science. *The 3rd ASE Big Data Science 2014 Beijing China*. Beijing.

Jie Cao , Zhan Bu, Guangliang Gao，Haicheng Tao, Weighted modularity optimization for crisp and fuzzy community detection in large-scale networks, *Physica A: Statistical Mechanics and its Applications*, 2016, 462（2016）:386-395.

Jie Cao, Zhiang Wu, Junjie Wu, Hui Xiong. SAIL: Summation-bAsed Incremental Learning Information-Theoretic Text Clustering. *IEEE Transactions on Cybernetics*. 2013, 43(2): 570-584.

Chong, A. Y., Li, B., Ngai, E. W., Eugene, C., & Lee, F. (2015). Predicting online product sales via online reviews, sentiments, and promotion strategies. *International Journal of Operations & Production Management,* 36 (4), 358-383.

Chong, A., Ch'ng, E., Liu, M., & Li, B. (2015, 6 24). Predicting consumer product demands via Big Data: the roles of online promotional marketing and online reviews. *International Journal of Production Research,* 1-15.

Christopher, S., Jeff, C., & Judy, S. (2016). Predicting social networking sites continuance intention through alternative services. *Industrial Management & Data Systems,* 117 (6), 1127-1144.

Conover, D. M., Gonçalves, B., & Flammini, A. (2012). Partisan asymmetries in online political activity. *EPJ Data Science,* 1 (1), 6.

Criado, J., Ignacio, Martínez-Fuentes, G., & Silván, A. (2012). Social media for political campaigning. The use of Twitter by Spanish Mayors in 2011 local elections. *Web 2.0 technologies and democratic governance*. 5, pp. 219-232. New York: Springer.

Daniel, G., Metaxas, P. T., & Mustafaraj, E. (2011). Limits of Electoral Predictions Using Twitter. *The 5th AAAI International Conference on Weblogs and Social Media (ICWSM 2011)* (July 17-21, 2011 Barcelona, Spain) .

David, L., Alex, P., & Lada, A. (2009). Computational Social Science. *Science*, 323 (5915), 721--723.

DipNote, B. (2009, 2 20). Retrieved from How Might the U.S. Utilize Innovative Technologies to Discuss U.S. Foreign Policy?: http://blogs.state.gov/index.php/entires/how_might_the_u.s._utilize_innovative_technologies_to_discuss_u.s._foreign_/

Dwi, P., Nugroho, & Hauff, C. (2015). Twitter-based Election Prediction in the Developing World. *Proceedings of the 26th ACM Conference on Hypertext and Social Media,* pp. 149-158.

Eugene, C. (2014). The Value of Using Big Data Technology in Computational Social Science. *The 3rd ASE Big Data Science 2014*, Beijing China. Beijing.

Giles, J. (2012). Computational social science: Making the links. *Nature*, 488, 448--450.

Goel, S., J., M. H., & Lahaie, S. (2010). Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (41), 17486-17490.

Gomez, J. (2014). Social Media Impact on Malaysia's 13th General Election . *Asia Pacific Media Educator* , 24 (1), 95-105.

Holotescu, C., G, G., & Bran, R. (2011). Microblogging meets politics. The influence of communication in 140 characters on Romanian presidential elections in 2009. *Revista Română de Comunicare şi Relaţii Publice* , 3 (13), 26-42.

I., K. (2009, 6 16). Daily Press Briefing #99. Retrieved from http://www.state.gov/r/pa/prs/dpb/2009/124991.htm

J., L., Goncalves, B., Ramasco, J., & Cattuto, C. (2012). Dynamical Classes of Collective Attention in Twitter. *Proceedings of the 21st International Conference on World Wide Web (WWW2012)* (April 16-20, 2012 Lyon, France), (pp. 251-260).

Johan, B., Mao, H., & Zeng, X. (2010). Twitter mood predicts the stock market. *Journal of Computational Science* , 2 (1), 1-8.

Kitsak, M., Gallos, L. K., Havlin, S., & Liljeros, F. (2010). Identification of influential spreaders in complex networks. *Nature physics*, 6 (11), 888--893.

Kwak, Haewoon, et al. "What is Twitter, a social network or a news media?" *Proceedings of the 19th international conference on World wide web*. ACM, 2010.

Lalitha, M. (2013). The Impact of Social Media in Social and Political Aspects in Malaysia: An Overview. *International Journal of Humanities and Social Science*, 3 (11), 71-76.

Lim, Y. (2013, 2 27). PM: GE13 will be Malaysia's 1st 'social media election. Retrieved from The Star Online: http://www.thestar.com.my/news/nation/2013/02/27/pm-ge13-will-be-malaysias-1st-social-media-election/

Macnamara, J., & Ansgar, Z. (2013). Social media communication in organizations: The challenges of balancing openness, strategy, and management. *International Journal of Strategic Communication* , 3 (11), 71-76.

Metaxas, Takis, P., & Mustafaraj, E. (2012). Social Media and the Elections. *Science* , 338 (6106), 472-473.

Muniandy, L. (2013). The Impact of Social Media in Social and Political Aspects in Malaysia: An Overview . *International Journal of Humanities and Social Science* , 3 (11), .

Ngai, E. W., Moon, K.-l. K., Lam, S. S., Chin, E. S., & Tao, S. S. (2015). Social media models, technologies, and applications. *Industrial Management & Data Systems* , 115 (5), 769-802.

Ostwald, K. (2013). How to win a lost election: Malapportionment and Malaysia's 2013 general election. *The Round Table* , 102 (6), 521--532.

Sitaram, A., & Huberman, B. A. (2010). Predicting the Future with Social Media. *Web Intelligence and Intelligent Agent Technology (WI-IAT)* (August 31, September 3, 2010 Toronto, Canada), (pp. 492-499). Toronto.

Sun, H.-l., Ch'ng, E., Yong, X., Garibaldi, J., & Chen, D.-b. (2018, April 15). A Fast Community Detection Method in Bipartite Networks by Distance Dynamics. *Physica A: Statistical Mechanics and its Applications* , 108-120.

Sun, H.-l., Ch'ng, E., Yong, X., Garibaldi, M. J., See, S., & Chen, D.-B. (2017). An improved game-theoretic approach to uncover overlapping communities. *International Journal of Modern Physics C* , 28 (8), 1750112-1750129.

Tumasjan, A., Sprenger, T. O., & Sandner, P. G. (2010, 10). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media* , pp. 178-185.

Tumasjan, A., T., O. S., & P., G. S. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *4th International AAAI Conference on Weblogs and Social Media (ICWSM)* (May 23-26, 2010 Washington). Washington.

Wong, F. M., Tan, C. W., Sen, S., & al., e. (2016). Quantifying political leaning from tweets, retweets, and retweeters. *IEEE transactions on knowledge and data engineering* , 28 (8), 2158-2172.

Table 1 Political parties and their representative Twitter accounts in MGE2013

Political Parties	Representative Twitter Accounts	Descriptions
BN	@NajibRazak	Najib is the President of leading ruling party BN.
	@barisanasional	It is the official Twitter account of BN.
	@DrRaisYatim	Dr Rais bin Yatim was a minister in multiple federal governments from 1974 to 2013.
	@Khairykj	Khairy Jamaluddin is the president of the youth wing of the United Malays National Organisation (UMNO).
PKR	@anwaribrahim	Anwar Ibrahim is the leader of PKR.
	@n_izzah	Nurul Izzah Anwar is the Vice President of PKR.
	@NajwanHalimi	Najwan Halimi is the former special officer of Anwar Ibrahim.
	@drwanazizah	Dr. Wan Azizah binti Wan Ismail is the leader of Opposition of Malaysia.
DAP	@limkitsiang	Lim Kit Siang is the prominent leader of DAP.
	@hannahyeoh	Hannah Yeoh is a member of DAP. She was reelected as Subang Jaya state assembly member in MGE2013.

	@cmlimguaneng	Lim Guan Eng is the Secretary General of DAP.
PAS	@ tuan_ibrahim	Tuan Ibrahim Tuan Man is the deputy president of PAS.
Mass Media	@staronline	The Star Online comes from the largest paid English newspaper The Star.
	@hmetromy	Harian Metro is Malaysia's first daily afternoon tabloid in Klang Valley.
	@tm_insider	The Malaysian Insider is a popular Malaysian bilingual news site.

Table 2 Topological features of propagation networks on the election day and the day after of the MGE2013.

Networks	#user	#retweet
Retweet network on **election** day	265,143	319,252
Retweet network after **election** day	95,201	86,567

Table 3 Community divisions of Twitter propagation network on election day of MGE2013. C indicates the clustering coefficients, l denotes the average shortest path length, and r represents the assortativity.

Political Parties	#Com	#Node	#Edge	C	l	r
BN	5	9745	22206	0.1037	3.956	-0.161
PKR	1	19817	37963	0.0311	4.565	-0.230
	14	2234	2370	0.0316	4.586	-0.339
DAP	3	12828	22166	0.0509	4.405	-0.129
PAS	27	325	354	0.0484	7.263	-0.219
Media	9	4578	5528	0.0258	5.759	-0.219
	4	11197	20262	0.0485	4.915	-0.188
	3	12828	22166	0.0509	4.405	-0.130

Table 4 Coreness scores of selected Twitter accounts on the election day of MGE2013. Core (Coreness equals 28), Transition (Coreness ranging from 21 to 27), Middle (Coreness ranging from 11 to 20) and Periphery (Coreness ranging from 1 to 10).

Political Parties	Representative Twitter Accounts	Coreness	Position
BN	@NajibRazak	28	Core
	@barisanasional	28	Core
	@DrRaisYatim	8	Periphery
	@Khairykj	4	Periphery
PKR	@anwaribrahim	23	Transition
	@n_izzah	10	Periphery
	@NajwanHalimi	12	Middle
	@drwanazizah	7	Periphery
DAP	@limkitsiang	12	Middle
	@hannahyeoh	11	Middle
	@cmlimguaneng	7	Periphery

PAS	@ tuan_ibrahim	4	Periphery
Media	@staronline	21	Transition
	@hmetromy	12	Periphery
	@tm_insider	12	Periphery

Table 5 Influential Spreaders and their tweets.

Twitter accounts	Supporting party	Sample tweets
P*2**	BN	Zairil stardom cant help him against Ng Song Kheng in Bukit Bendera BN is winning there comfortably! #PRU13 #GE13 #UbahPenang
K*s**	BN	These are clear signs of support for @NajibRazak #Vote4BN #NajibMyPM #PRU13 #GE13 #JomUndi.
H*t**	PR	PR has won We urge UMNO and the EC to not attempt to hijack the results. #ubah #GE13
M*a**	-	The victor or loser of this 13th General Election will not be BN or PR It will be Malaysia its people and our children.
U*u**	-	My friend is very happy that now he can apply 2 buy a house under PR1MA scheme! good 4 him! #GE13 #PRU13
K*N**	PR	God willing PR can win. Kita dah usaha yg lain kita. (God willing PR can win. We and our other efforts.)
D*i**	-	Its a very novel way of helping 2 ease people burden of rising living costs by abolishing school fees in govt school!
T*l**	BN	The BBC reports most Malaysians are satisfied with the direction of their country #BetterNation Vote @barisanasiona.
Z*n**	PR	Early Congratulation to Dato' Seri Anwar Ibrahim and PKR.
A*t**	BN	People have decided - Saya pilih Najib Razak. (People have decided - I select Najib Razak.)
Z*i**	PR	My petrol tank is empty but it's alright I know we have good news tmr #ge13 @anwaribrahim
H*a**	BN	Good Morning @barisanasional ! I'll vote for you and @najibrazak forever #tolakPR #ge13
A*t**	BN	@NajibRazak good luck PM & all the BN candidates Our future lies with you.

Table 6 Community divisions of Twitter propagation network on the election day of MGE2013. C indicates the clustering coefficients, l denotes the average shortest path length, and r represents the assortativity.

Political Parties	#Com	#Node	#Edge	C	l	r
BN	2	3797	5427	0.1049	3.918	-0.337
PKR	5	9745	22206	0.1037	3.956	-0.161
DAP	4	3353	4460	0.0242	5.516	-0.207
	6	2230	2527	0.0191	4.417	-0.258
PAS	2	-	-	-	-	-
Media	4	3353	4460	0.0242	5.516	-0.207
	31	221	225	0.0200	5.956	-0.318

Table 7 Coreness scores of selected Twitter accounts after the election **day of MGE2013. Core (Coreness equals 26), Transition (Coreness ranging rom 21 to 25), Middle (Coreness ranging rom 11 to 20) and Periphery (Coreness ranging rom 1 to 10).**

Political Parties	Representative Twitter Accounts	Coreness	Position
BN	@NajibRazak	26	Core
	@barisanasional	26	Core
	@DrRaisYatim	8	Periphery
	@Khairykj	2	Periphery
PKR	@anwaribrahim	7	Periphery
	@n_izzah	6	Periphery
	@NajwanHalimi	3	Periphery
	@drwanazizah	7	Periphery
DAP	@limkitsiang	6	Periphery
	@hannahyeoh	3	Periphery
	@cmlimguaneng	3	Periphery
PAS	@ tuan_ibrahim	1	Periphery
Media	@staronline	6	Periphery

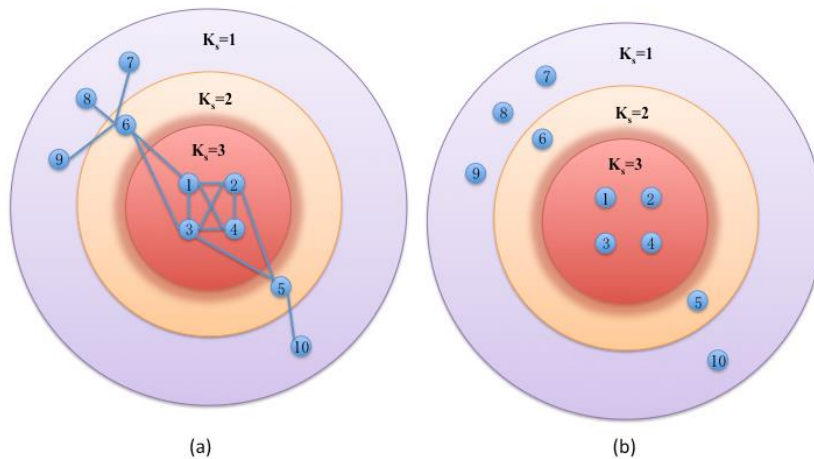|		
@hmetromy	6	Periphery
@tm_insider	3	Periphery

Figure 1 A dummy sample of k-shell decomposition applied to a network with 10 nodes. (a) Node 1, 2, 3, 4 locate at the dark red core zone with k_s=3, node 5 and 6 have k_s=2 and node 7, 8, 9, 10 have k_s=1. (b) The network topological structure is ignored and nodes are assigned to different layers according to their k_s scores.
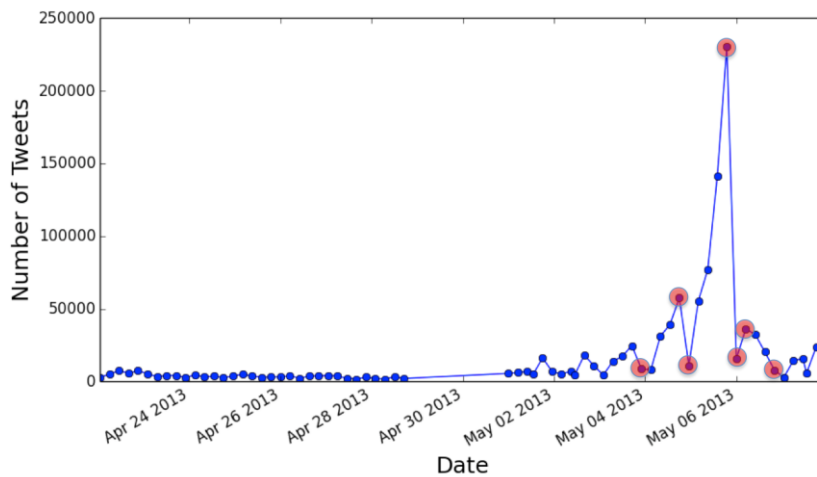

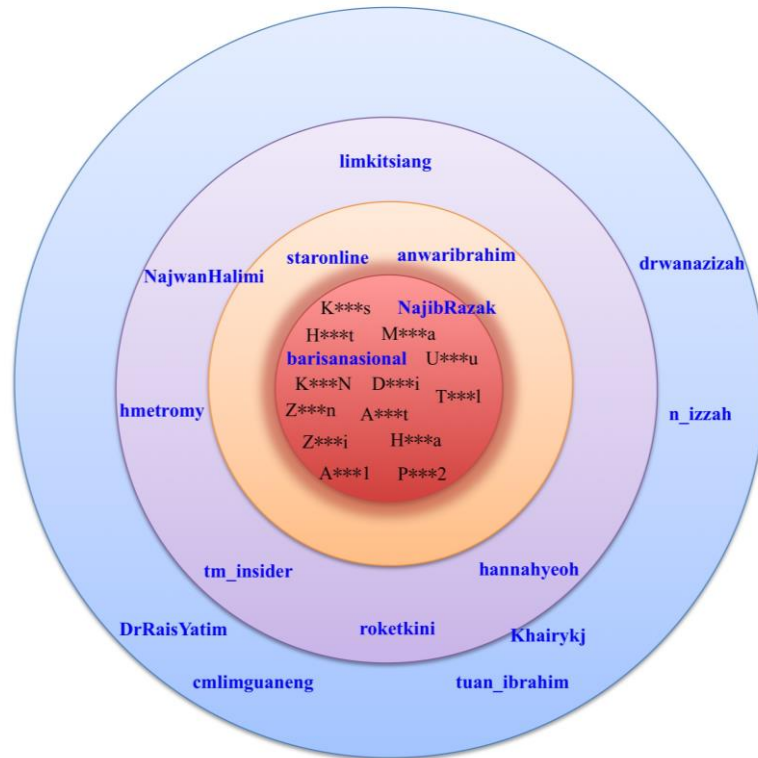**Figure 2 Number of Tweets every five hours during MGE2013**

Figure 3 The Twitter propagation sphere on the election day of MGE2013. Personal Twitter account names have been made anonymous to protect privacy, but public figures remain open to public.
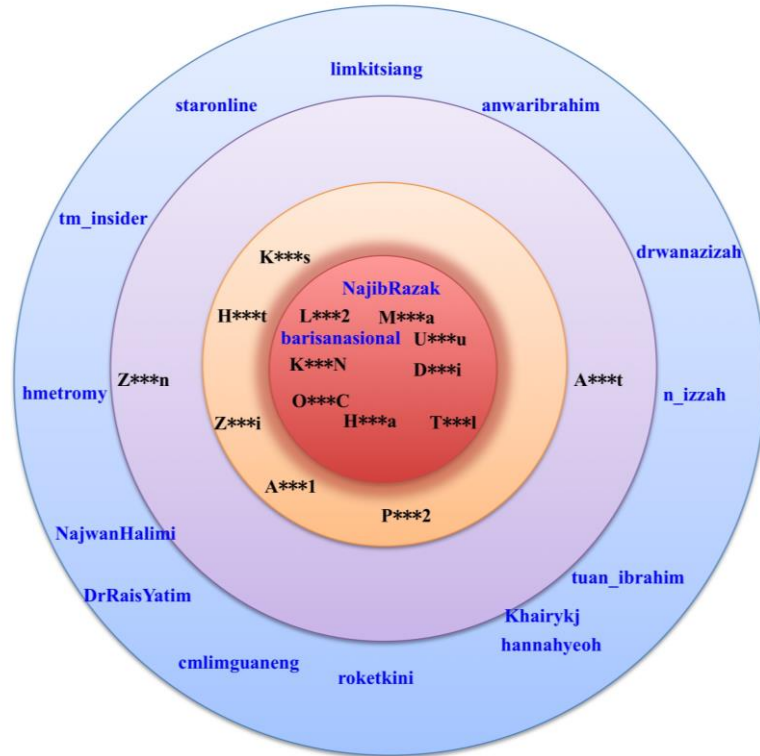
Figure 4 Tweets propagation sphere after the election day of MGE2013