Some Inquiries to Spontaneous Opinions: A case with Twitter in Indonesia

Maulana, Ardian and Situngkir, Hokky

Bandung Fe Institute

30 October 2010

Some Inquiries to Spontaneous Opinions

A case with Twitter in Indonesia

Ardian Maulana
[ai@compsoc.bandungfe.net]
Dept.Computational Sociology
Bandung Fe Institute

Hokky Situngkir
[hs@compsoc.bandungfe.net]
Dept.Computational Sociology
Bandung Fe Institute

Abstract

The paper discusses opportunities to utilize the series of micro-blogs as provided by the Twitter in observation of opinion dynamics. The spontaneity of tweets is more, as the service is attached more to the mobile communications. The extraction of information in the series of tweets is demonstrated as in conceptual map and mention map. From the latter, the social network stylized properties, *i.e.:* power law distribution is shown. The exemplification of the methodology is on the 82nd commemoration of Indonesian Youth Pledge and the participatory movement of Indonesian capitol city, Jakarta.

Keywords: Twitter, social network, social media, text analysis, conceptual map, mention map.

Life is a series of natural and spontaneous changes.
-Lao Tzu-

We came across the word "twitter," and it was just perfect. The definition was "a short burst of inconsequential information," and "chirps from birds." That's exactly what the product was.
-Jack Dorsey-

1. Introduction

Blogging, as well as micro-blogging, is no question to play important roles in the process of modern democracy. The existence of internet social network applications, *e.g.:* Facebook, MySpace, Friendster and the internet social media *e.g.:* Twitter, Plurk, whatsoever have shares in shaping of social and political life in countries around the world. The growing size of accounts in Indonesia reflects this. Evans (2010) reported that from a survey of 13 million Twitter accounts, the fifth largest countries in the world using Twitter is Indonesia – holding 2.34% of the sampled twits – after United States, United Kingdom, Brazil, Canada, and Australia.

Twitter is a microblogging service in which users (called *tweeps*) post their status (called *tweets*) about anything in 140 characters. One can follow other users to receive their tweets, but uniquely with other online social networking services, the user being followers need not to follow back. Thus, a user with large amount of followers is more likely to "broadcast" her message to all her followers. Twitter is more likely to be the synthesis of chat rooms and web-blogs, because a user can respond to other user through her tweets. There four unique ways to post messages in Twitter, i.e.:

a. A retweet: a user receiving tweet from one she is following and repost it to her own followers. Most of twitter clients sign this kind of posting with the code "RT" and then followed by the account name she is retweet-ing. An account name can be included in a tweet by the sign "@". This mechanism is actually the thing give Twitter a strong media for spreading information from persons to persons.
b. A reply: a user can respond a tweet she is receiving from whom she is following. This type, in fact, can be a way that turning Twitter to be a sort of chat service. A reply will be signed with the sign "@" (and followed by the one whose tweet she want to reply) at the very beginning of the posting.
c. A mention: user post a tweet by mentioning other user. In this case, a user included the sign "@" (and followed by the name of account she want to mention) in the middle of the message.
d. A hashtag: user post a tweet begun with the sign "#" and then followed by certain keywords. Other users can search any of the hashtag later on.

Twitter becomes widely recognized for its open API (application programming interface) that would eventually makes it easily to be adaptable to wide ranges of devices and operating platforms available in the market. Interestingly, the growing mobile communication services worldwide, and consequently, Indonesian mobile-phone market, has made Twitter one of favorites in mobile internetworking. Thus, people are sending (or broadcasting – for one followed by great numbers of tweeps) information just like texting using the short message services (which is also way to post tweets that is also available in some countries without any client software).

So when it comes to social issues, rumors, gossips, and even news, Twitter has made it possible anyone post any fast and spontaneous expression or comment by a slight retweet, or a mention, or a reply, or making community conformed keywords related to things by employing hastagged tweets through computers and the exponentially growing mobile devices. A report from Spire Research (2009) reveals the rapid growing rate of mobile internet user in Indonesia, from 34% (2008) to 44% (2009). The launch and marketing of Blackberry Services in the country since 1999 has made a great impact on the numbers substantially since the services offer lower charged internet through mobile devices as provided in special packages by Indonesia's cellular providers, the Indosat, Telkomsel, Excelcomindo, *et cetera* – the market shares among them are shown in figure 1.
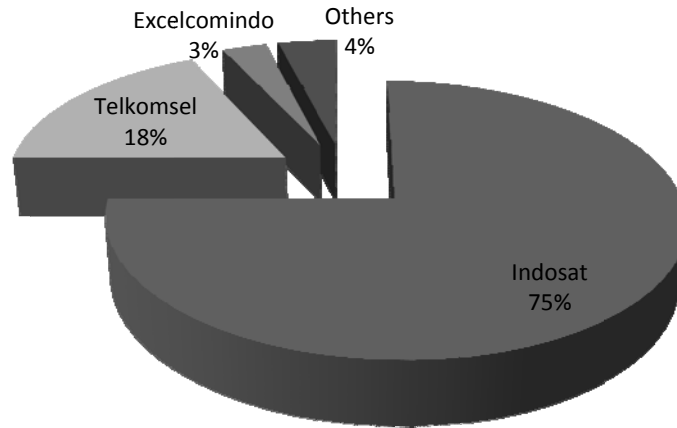
Figure 1
The market-share of mobile internet providers in Indonesia[*]

More than 80% handhelds in Indonesia is internet-enabled, 53% of the users are between the ages of 18 and 27, and 82% of the total number of mobile web users are male[*]. From January 2009 to June 2009, mobile ad requests on Blackberry phones increased by 842% and the strong competitors iPhones which increased by 205%. These numbers are positively correlated with the growing numbers of Twitter users in Indonesia as previously shown.

2. Tweets and inquiries to opinions dynamics
Twitter has been becoming a kind of opinion reservoirs related to all things to be discussed and diffused issues online. It contains all opinions about all things in the world channeled through the internet. In-depth look into stylized properties of Twitter as reported by Cheng, *et. al.* (2009) shows that Twitter is more likely to be a (internet based) social media rather than social networking service. The topological analysis of the follower-following users of Twitter is confirmed a deviation to the regular social network as shown previously by Newman & Park (2003). Twitter's topology is non-power-law follower distribution, short effective diameter, and very low reciprocity (Kwak *et. al.*, 2010). Twitter is more likely to be an extension to the conventional mass media, as more than observed 85% of the global trending topics in it are becoming headline news.

---

[*] As broadcasted by InMobi URL: http://blackberryrocks.com/2009/09/30/statistics-blackberry-beating-iphone-indonesia/

Thus, in the global perspective, Twitter can be seen as the record of accumulation what are people saying over time. Spontaneity of people using their mobile phone "randomly" looking at their micro-lives as well as the global issue. If one day, we might have a tank filled with all the information reflecting what people are having in life, it should have been just like (or be) Twitter. As reported by the National Public Radio (2010), the Library of Congress of the United States of America is archiving the Twitter.
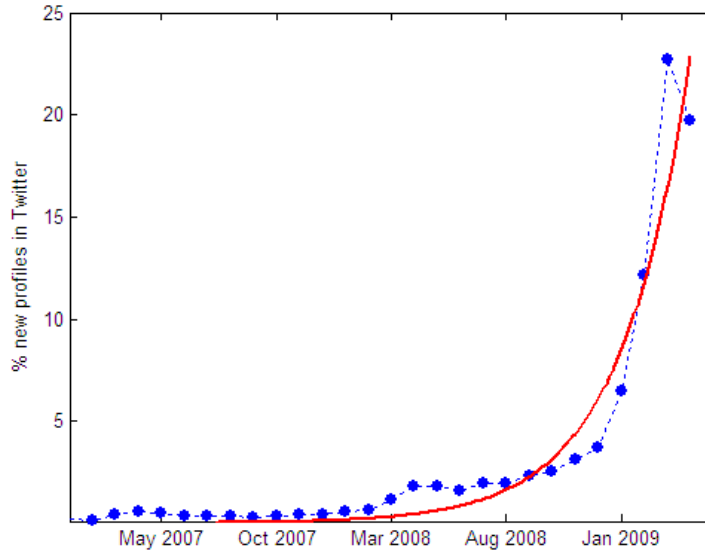
Figure 2
The exponential growth of the population in Twitter

The task of this report is to propose an alternative to capture information from the population of tweets and tweeps. From the previous section, we can see that observation to twitter can be conducted tweep-wise and tweet-wise. The tweep-wise analyzes the behavior of the Twitter users and the tweet-wise seeing the population of tweets as (textual) corpus.

A tweet $s_x \in S, x = \{1,2,3,...\}$ are possibly consisted of words (represented as concepts), $k \in K$ and mentions, $m \in M$. Thus by letting words to be represented as concepts (excluded the pronouns, adjectives, adverb, verb, and other grammar sensitive ones) and denoting them as fully connected graph, the series of tweets are thus being treated as a conceptual graph, $\Omega^K = \left(K, E^k \right)$ and user-network based on mention (mention map), $\Omega^M = \left(M, E^m \right)$ where $E^k = \left\{ e_{ij}^k \right\}$ as connection between two concepts, and $E^m = \left\{ e_{ij}^m \right\}$ as connection between two user accounts, as m_i mention m_j in the particular tweet. Thus, in this analysis we propose two models of graph in our observation to the population of tweets: conceptual graph and the mention map.

In conceptual graph, $\Omega^K = \left(K, E^k \right)$, if two concepts k_i and k_j are used once in a tweet, then the value $e_{ij} = 1$. Thus. Our conceptual graph would be a weighted graph where the total strength of the edge is written,

$$< e_{ij} >= \sum_{i,j}^{N} e_{ij} \qquad\qquad (1)$$

and

$$< h_i >= \sum_{j}^{N} e_{ij} \qquad\qquad (2)$$

where $< e_{ij} >$ and $< h_i >$ both represent the total number of the usages in the corpus. The variable $< h_i >$ shows specifically how many the particular word k_i used in the whole tweet population.

The similar model is also implemented or the mention map. We define the network as a directed weighted graph, $\Omega^M = \left(M, E^m \right)$, where each vertex in $m_i \in M$ is twitter user who engaged in conversation/discussion. For any two users, m_i and $m_j \in M$ there is an arch from m_i and m_j if and only if user m_i mentions or re-tweets m_j's tweet. In the other words, we introduce the arch from m_i to m_j as long as m_i posts a tweet that contains m_j's user name, while the weight of the arch equivalences to a total number of the message similarly to eq. (1). There are probably more than one arch created from a message because a user often mentions more than one user in it.

The two drawings give two different kinds of information from two networks. The first is how concepts relate to each other and the second one how someone related to one another in the online in the sphere of discourses.

3. Cases of Study
In order to understand how the proposal might be useful to the further information acquisitions textual and micro-blogging services like Twitter, we will have two cases for further study. The first is a sort of poll initiated by some nationally recognized tweeps. The second, we observe the dynamics activism initiated in Twitter regarding to particular political issue.

   a. **"Poll" with Twitter**
 A very important historical events in Indonesia is the Youth Pledge (*Sumpah Pemuda*), a declaration made by young Indonesian from many places in the archipelago in a rally for movement against the colonialism by the Dutch East Indies, October 28th 1928. As we know, the diversity of cultural aspects for the large archipelagic geography has made the movement previously to be sporadic, and thus felt to be very ineffective (Kahlin, 1952). The differences of ethnicities, languages, and ideal paradigms for freedom have been very high. Yet, the young people proclaimed the well-known three ideals: one motherland, one nation and one language. From the historical milestones, the pledge was a breakthrough: a unified recognition of the diversity in Indonesia.
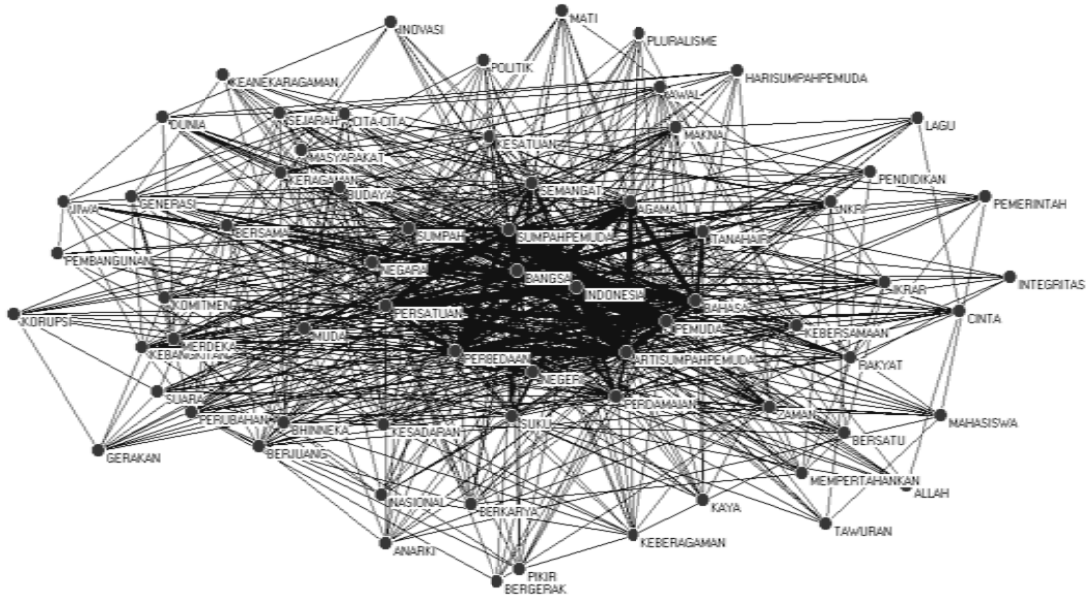
In the commemoration of the Youth Pledge, a twitter activist, Fahira Idris[**] create a "free association" poll with hashtag #28OktPoll[***] asking her thousands of followers what they might think in their actual daily life of the Youth Pledge. More than 4000 tweets are gathered. A question might rise, how can we read their answers and possibly have the summary instantly? The nature of hashtags and mentions in Twitter has made reading tweets not enjoyable. Thus, the analysis and

---

[**] http://www.twitter.com/fahiraidris
[***] http://search.twitter.com/search?q=%2328OktPoll

data representation as described in the previous section is brought. The result is a interconnected concepts relating to it, and become interestingly easy to read and summarize. As shown in figure 3, the concepts surrounding the theme of the "commemoration to Indonesian Youth Pledge" are looked orbiting.

Figure 3
The network reflecting how Indonesian Twitter Users
recognize commemoration of Indonesian Youth Pledge after 82 years.

Indonesian Youth Pledge 1928 is a celebration of peaceful declaration for unity in diversity. Thus some concepts related to endeavor to left differences and focus on the collective aims for independence instead. While tweeps posted their feeling about the commemoration, some actual concepts arise.

While the picture has given some clues to the actual concepts, a centrality measures of weighted graph may also be calculated for more comprehensive answer. As it has been shown in Situngkir (2007). Centrality measures reveal the most important concepts in such swarming concepts in the network. One of the large amounts of the centrality measures are the eigen centrality, h_i, which is made based on an assumption that a concept should be important as it has been related with other important one. An algorithmic can be applied (as to be used by the Google search engine), conjectured by Brin & Page (1998).

The centrality measures based on the centrality eigen can be written mathematically,

$$c_i = \left(\frac{1}{\lambda}\right)\sum_j E^k c_j \qquad (3)$$

where λ is a constant and E^k the adjacency matrix of the interconnection among concepts. It can also be written as

$$\lambda \, \overline{c} = E \, \overline{c} \qquad (4)$$

where $\overline{c}$ is the eigen vector of the rexpective adjacency matrix.

Table 1
Rank of most important concept based on eigen centrality measures in #28

Rank of Centrality	Concept	in English
1	*'SUMPAHPEMUDA'*	**Youth Pledge**
2	*'INDONESIA'*	**Indonesia**
3	*'PEMUDA'*	**Youth**
4	*'BANGSA'*	**Nation**
5	*'ARTISUMPAHPEMUDA'*	**Meaning of Youth Pledge**
6	*'PERBEDAAN'*	**Differences**
7	*'PERSATUAN'*	**National Unity**
8	*'BAHASA'*	**Languages**
9	*'AGAMA'*	**Religions**
10	*'SUMPAH'*	**Pledges**

Table 1 shows the ten most important concepts as calculated with eq. (4). As we have been discussed before, the Indonesian Youth Pledge (*Sumpah Pemuda*) is a historical moments. Thus, we have guessed easily that most of the important concepts would be related to the social and political problems at the October 28th 1928. However, an actual concept that was not at all related to the criticalities 82 years before is having a big centrality measure. The discussion about religious fanaticism that frequently threatened the spirit of the national unity in actual daily life in Indonesia emerges. This thing is rarely related to the historical moments of the Youth Pledge 1928. Thus, the free association "poll" for actual impression here has pointed the a social problem that is actually detached from the events 82 years before. This can be easily continued to the concepts with lower centrality measures.

b. In search for activism for better Indonesian capitol

In other occasion, a hashtag #SaveJkt[****] was initiated by some social activist in the country. The hashtag was dedicated to the participation of civil society for the betterment ecological and social problems in the city, e.g.: flood during the rainy season, depressing traffic jam, and so on. Thus, this problem is naturally more complicated than our previous case of study. Twitter activists who come the surface as organization called "savejkt" created this hashtag.

For a period of six days starting from November 21st, 2010 to November 26th, 2010, we archived all public tweets that containing the hash tag #savejkt. There are total of 2466 post from 804 distinct users in this collection. We construct the mention Map of user interaction based on this data and obtain the graph of largest connected component comprise with 818 vertices and 2495 edges (Figure 4). The number of vertices in the graph is more than the number of users that posts the messages. This is because of we count all users that is mentioned on a messages although they never post any messages containing hash tag #savejkt.

[****] http://search.twitter.com/search?q=%23SaveJkt

Like we have conducted for the hashtag #28OktPoll, we study conversation activities among twitter members whose tweets containing the hashtag #savejkt. Consequently to the nature of the theme, this hashtag is a kind of tweet's topic related to Jakarta city and its problem.

Figure 4
The mention map of user interaction on hashtag #savejkt

Here we provide statistical properties of the collected data. Figure 5 shows the cumulative distribution function of tweets containing hash tag #savejkt plotted on the log-log scale. One can observe that the number of users has tweets larger than *x* is an inverse power of *x*, as follow:

$$P(X \geq x) \propto x^{(-\alpha+1)} \tag{5}$$

where α as the scaling parameter has value *2.11*. This pattern states few tweeps post tweets disproportionately more times than the majority of users.

When we detail our observation to the mention map representing the user interaction network, we discover that the network also exhibits power law pattern in their strength distribution. Here we define the strength of vertex as a natural generalization of vertex degree to weighted graph, as follows:

$$s_i = \sum_{j \in v_i} w_{ij} \tag{6}$$

and

$$s_j = \sum_{i \in v_j} w_{ji} \tag{7}$$

where s_i and s_j are out-strength and in-strength of vertex.

Figure 6 shows that the mention graph is a scale-free network, which is characterized by power law distributions of strength connectivity with scaling parameter of *2.05* for in-strength distribution and *2.07* for out-strength distribution. This fact states few vertices that are very connected, but most are weak ones. The highly concentration of in-strength connection in small fraction of tweeps means that those users become the authoritative source of information which is referred by majority of users. Likewise, the scaling pattern in out-strength distribution implies that a few users are very

communicative and responsible for a disproportionately high fraction of discourse related the hash tag of #savejkt, as shown in figure 5. The scale-free characteristics of the mention graph has also been reported by Ediger et.al (2010) in terms of degree distribution.
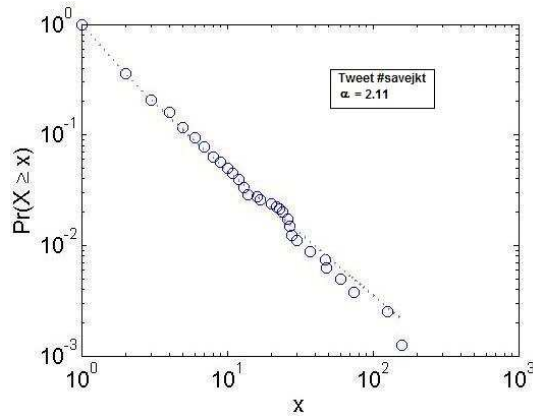
Figure 5
Cumulative distribution function of tweet witihin tag of #savejkt

Power-law distributions happen in an extraordinarily diverse range of phenomena (Newman, 2004). The emergence of scaling phenomena in the interactions and communications among Twitter users indicates that this online community also shares the statistical characteristics of the complex system. This finding implies that human interactions, which emerge those scaling law, cannot be treated as a kind of random process. However, this is opens a question of the origin of mechanism responsible for the emergence of scaling pattern in the twitter user communication. Thus, referring to the work of Kwak *et. al.* (2010), while the follower-following in Twitter is not at all reflecting the social network, when it comes to the interaction inside of it, the stylized social network exhibits.
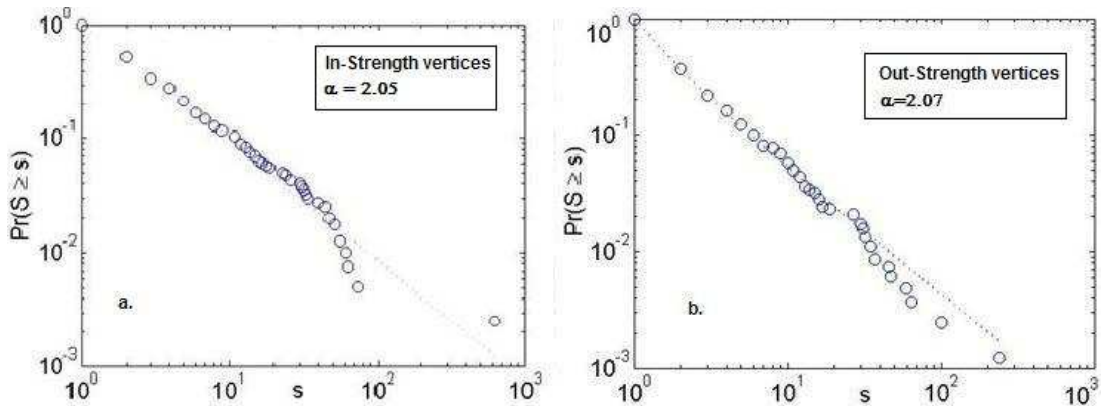
Figure 6
Cumulative distribution function of strength distribution: a. In strength ; b. Out strength

We continue our work by examining to see the importance of tweeps in the series of the tweets related to the hashtag as shon in the Mention Map. Regarding the nature of twitter as social media

Kwak, *et.al.,* (2010), we think that the notion of importance should be related with concept of authority in term of information sources. The more other users mention a user, the more important the user is. However, the quality of others that mention him should contribute to his importance. In the other words, the importance of a user depends on the number as well as the importance of other users mentioning him. By incorporating this assumption, again we use the eigenvector centrality measure as discussed previously through eq. (4).

Table 2
Top 20 Users by Eigen centrality

Rank	Twitter Account	Profession	Position in "SaveJakarta" Organization
1	savejkt		Site account
2	billytanjung	Praktisi Media Online	
3	yanuarnugroho	Researcher	Strategy Team
4	anggaadoet	Lecturer	Concept Team
5	obsat	Site Account of "Obrolan Langsat"	
6	andibachtiar	Film Director	Campaign Team
7	Reiza_Patters	Activist of Environmental Issues	Campaign Team
8	unilubis	Chief Editor News ANTV	
9	iwibisono	Activist of "Green Aceh" Organization	
10	deekasuksmana	Practitioner	
11	enda	Activist of Social Media	Campaign Team
12	rinohermadi	Practitioner	
13	sociotalker	Lecturer	Strategy Team
14	IndraJPiliang	Politicians	
15	abah__	Web Designer	Campaign Team
16	bangwinissimo	Activist of Social Media	Campaign Team
17	cheestea	Interior Designer	Concept Team
18	PutraJayaHusin	Politicians	
19	voiceofjakarta	Site account of "Voice Of Jakarta"	
20	yunartowijaya	Researcher	Strategy Team

The result is interesting as shown in Table 2. It shows the highest ranked users in the mention map. By identifying the detailed information of twitter users who are listed in the table, we find that the top ranked user are dominated by the initiator of "savejkt" organization and eventually emerging the new recruit to the discourse, regarding to people that latter are playing important roles in this #SaveJkt hashtag.

4. Concluding Remarks
While microblogging has come to the tendency for spontaneous mobile communication, growing wider, and faster micro-social response, some conjectures for acquisition to extracting and thus grasping more information are in need. Twitter is constituted concepts and mentions to other people in large scalability.

When our focus is the message, observing the conceptual map is brought as the mention map, when the emphasis is the people. The textual analysis brought by the conceptual map may reveal some important concepts that cannot be easily seen just by reading the tweets. The concept of centrality measures help to do this.

Twitter as a social media service has unique properties exhibiting interesting stylized facts. Our observation discusses that when the focus is the way users mentioning others, a stylized fact about social network is shown, *i.e.:* the power law distribution. The observation to the mention map give us the reflection on who is being dominant and has important position as discussions go.

Acknowledgement

Works Cited:

Brin, S. & Page, L. (1998) "The Anatomy of a Large-Scale Hypertextual Web Search Engine". *Computer Networks* 30: 107-17.

Cheng, A., Evans, M., & Singh, H. (2009). *Inside Twitter: An In-Depth Look Inside the Twitter World*. Sysomos Inc. URL: `http://www.sysomos.com/insidetwitter/`

Ediger, D., Jiang, K., Riedy, E.J., Bader, D.A., Corley, C.D., Farber, R.M., & Reynolds. W. (2010). "Massive Social Network Analysis: Mining Twitter for Social Good. In *Proceedings of the International Conference on Parallel Processing*. San Diego, CA 2010

Evans, M. (2010). *The Top Twitter Countries and Cities. Sysomos*. On-line publication: `http://blog.sysomos.com/2010/01/14/exploring-the-use-of-twitter-around-the-world/`

Kahin, G. M. (1952). *Nationalism and Revolution in Indonesia*. Cornell UP.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). "What is Twitter, a social network or a news media?". *Proceedings of the 19th international conference on World wide web* pp. 591-600. ACM.

National Public Radio. (2010). *Library Of Congress Is Archiving Your Tweets*. URL: `http://www.npr.org/templates/story/story.php?storyId=126086325`

Newman, M. E. J.(2004)."Power laws, Pareto distributions and Zipf's law". *Contemporary Physics* 46 (5): 323-51.

Newman, M. E. J. & Park, J. "Why social networks are different from other types of network." *Physical Review E* 68(3).

Situngkir, H. (2007). "Konsep Sentralitas dalam Jaringan Teks". *BFI Working Paper Series* WPV2007. Bandung Fe Institute.

Spire Research. (2009). *Indonesia: Asia's mobile internet success story*. Spire Research & Consulting. URL: `http://www.spireresearch.com/pdf/archive/ejournal-dec09/Indonesia-Asia's mobile internet success story.pdf`