

# The Impact of Influential Leaders in the Formation and Development of Social Networks

Boram Park \*  
SK Planet  
45 4-ga Namdaemun-ro,  
Jung-gu  
Seoul, Korea  
mcboram@sk.com

Kibeom Lee  
Department of  
Transdisciplinary Studies  
Seoul National University,  
Seoul, Korea  
kiblee@snu.ac.kr

Namjun Kang  
Department of  
Transdisciplinary Studies  
Seoul National University,  
Seoul, Korea  
nj kang@snu.ac.kr

## ABSTRACT

This study investigates the communication patterns and network structure of influential opinion leaders on Twitter during the 2011 Seoul mayoral elections. Among the two candidates, we focus on the usage pattern of Wonsoon Park, who actively used Twitter during the election campaign. We analyzed the network structure of candidate Park and his 15 Twitter mentors during the election period (September 26, 2011 - October 26, 2011). The gathered data consists of 19,227 tweets from 8,547 users who were responded to by one of the 17 selected opinion leaders through mentions (@) or retweets (RT). To find the authorities and hubs, which play a crucial role in information propagation, the HITS algorithm was used to quantify the influence exerted by the opinion leaders. In addition, social network triads were used to identify the communication patterns between individual users on Twitter. Results of the analysis showed that the structure of the communication patterns in Twitter were mostly fragmented rather than transitive. This signified that communication occurred from, or converged to, a single node, rather than circulating through multiple nodes during the election period. The majority of the network structures were fragmented, or one-way conversations. In other words, communication happened in the form of aggregation and propagation, rather than sharing and circulating various ideas.

## Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: Sociology

## General Terms

Experimentation

\*Research on this study was done when author was affiliated with the Department of Transdisciplinary Studies, Seoul National University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*International Conference on Communities & Technologies 2013 Munich, Germany*

Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...\$15.00.

## Keywords

social network services; influence; twitter communication; virtual community;

## 1. INTRODUCTION

Society is facing various changes as the era of a new media, social network services, is opening. Now social network services, which are representatives of virtual communities, are impacting the way people communicate by crossing online and offline boundaries. Particularly, individuals who are active in social network services as opinion leaders are having a large influence in the formation and development of virtual communities. The importance of individuals, who exercise their influence by producing and transferring information, is surfacing again in the SNS environment. In a social network service, nodes of users are connected to a myriad of other users, forming and developing a network. This network creates various social groups in the SNS, which form different types of communities and communication compared to offline social groups.

Social network services (SNSs) are an emerging form of communication, which play a fundamental role as a medium for the spread of information, news, and influence among their users. One of the most notable SNSs is Twitter. Among the two billion Internet users in the world, Twitter had two hundred million users after its launch in 2006 and is still growing rapidly. The relationships between Twitter users are not bidirectional, which means that information flows asymmetrically, making people participate at different levels. The top 10% of prolific Twitter users accounted for over 90% of the traffic [2] and the top 0.05% of elite users accounted for almost half of all posted URLs in Twitter [11]. This implies that tweets, posts composed of 140 characters in Twitter, are produced and communicated by a few leading users with a one-way, one-to-many communication pattern. A few leading users in Twitter take the role of opinion leaders interpreting and disseminating information to the public. This study seeks to identify unique network structures and communication patterns in virtual communities, using Twitter as a representative of such communities. In particular, we measure the influence that opinion leaders have on the public and the entire virtual community in the social network service.

In this paper, we try to measure the influence of users who are representative of opinion leaders in the Twitter space during the 2011 Seoul mayoral elections. Opinion

leaders are defined as those who bring in new information, ideas, and opinions, then disseminate them down to the masses, thereby influencing the opinions and decisions of others through bi-directional communication ([7]). Opinion leaders capture the most representative opinions in the social network. They can play a key role in affecting decision-making through mutual interactions with the public. They are, therefore, important for understanding the massive and complex public and the public opinion [7]. This paper examines these influential individuals on Twitter through network structure analysis. We also analyze their structure, examine how they are connected to others, and assess whether they are at a position of influencing others. Among the two candidates that ran for the Seoul mayoral elections, we focus on Wonsoon Park, who selected his own Twitter mentors and actively used them during the election campaign, while using analysis results of candidate Na for basic data.

## 2. RELATED WORK

As a new developing medium, Twitter has been evaluated as a space in which people can freely communicate under an SNS environment. In this virtual space, individuals produce and disseminate information and start to have influence and power over others. Opinion leaders on Twitter have distinct characteristics, such as an increased network size, spreading of information via retweets, and the usage of common channels by communicators and intermediaries. These characteristics enable powerful individuals to have greater influences on many more people, so it is important to trace the individuals who have influence for understanding the formation and development of virtual communities.

With the gaining popularity of Twitter, a number of recent studies have examined the influence and information propagation in the Twitter space. Kwak et al. researched the follower-following topology of Twitter and its power as a new medium of information sharing [4]. They concluded that Twitter's highly skewed distribution and low reciprocity of followers closely resembled an information-sharing network than a social network. Kwak et al. classified the trending topics and found that the majority (over 85%) of topics were headline news or persistent news in nature. They compared three measures of influence: number of followers, PageRank, and number of retweets. The study found that rankings of the influences of users measured by number of followers and PageRank were both very similar. The rankings that were done using the number of retweets, however, did not show any similarities.

The study by Kwak et al. is meaningful in that it measured information propagation and ranked influences of users through followers, retweets, and the PageRank algorithm. Also, it has significance in that it is the first quantitative study on the entire Twittersphere and the information diffusion in it.

Weng et al. used and compared the number of followers and the PageRank algorithm to identify influential users on Twitter [10]. PageRank can measure the influence taking both the topical similarity between users and the link structures. They observed that (1) 72.4% of the users in Twitter follow more than 80% of their followers, and (2) 80.5% of the users have 80% of users they are following follow them back. They found that this was caused from the phenomenon of homophily. Their study is interesting in that it defined influential people with topics. In addition, it also measured

influence through the PageRank algorithm, ultimately quantifying and ranking users' influences.

Both of the above-mentioned studies went through the process of analyzing influence and information propagation on Twitter using the PageRank algorithm. Kwak measured influence using retweets as a metric, while Weng used topics as the measurement unit. In this study, influential users are analyzed using the HITS algorithm, as the algorithm is able to assign authority and hub scores for every node in a network. In addition, unlike the other studies, this research interprets users' influences based on traditional communication theory. Although the analysis on the data is done with algorithms, the interpretation of why such results came to be is done with communication tools from the social sciences.

Duncan J. Watts mainly researched network structures and individuals influences. In order to describe how individuals influence each other decisions, Watts specified who influences whom in a random network [9]. The study by Watts embodies the features of interpersonal influence networks. In particular, it has significance in that the study applied the traditional two-step flow of communication model, in which media, the source of information, played an important role on random networks. The influential individuals, instead of the media, is at the heart of the dissemination and convergence of information in Watts' influence network model.

Watts defined these kinds of individuals who have extreme influences on others as hyperinfluentials [9]. The most influential individuals in high-variance networks have roughly 40 times more influence than that in low-variance networks of the same average density. Thus, the presence of hyperinfluentials does indeed affect the size and prevalence of influence.

The influential hypothesis by Watts defined that large cascades of influence are driven not by influentials but by a critical mass of easily influenced individuals. Thus, influentials are less important than is generally supposed, either as initiators of large cascades or as early adopters ([9]). Watts' study is significant in that it analyzed Lazarsfeld's traditional theories [5] with large amounts of data and concluded that an individual's influential powers were less than previously thought. However, there is a need to verify whether individuals' influences are not effective in real-world events, such as elections, which are dealt with in this current study.

In addition, in the case of Korea, Twitter is used mainly by the younger generation, people in their 20s and 30s, and acts as an advocate of progressivism. In this special situation, it is expected that an individual's role in the new social group, Twitter, will have great influences contrary to the claims of Watts. Therefore, the current study examines the roles that a new media, such as Twitter, took in a real-world event (i.e. the elections). This is done through analyzing the communication patterns of the users and also analyzing the network structure.

Wu's 2011 study analyzes the communication relationships and the homogeneity of users on Twitter. Among the users, 'elite' users who attracted more than 50% of the attraction on Twitter were classified into four equal-sized categories: celebrity, media, organization, and blog. After analyzing who listens to whom by collecting the tweets that were exchanged between the elite categories, the study found that, save the blog category, the users of the remaining three categories interacted most frequently with groups that were homogeneous.

In the case of retweets, the influence patterns of the elite users were strongly homophilous. Analysis results showed that there were 93 retweets per elite user compared to 1.1 retweets per ordinary person in blogs. This can be seen as reflecting blogs’ characteristics of recycling and filtering. Wu’s study classified the top 0.05% influential elite users on Twitter. It is an interesting study in that it analyzes the influences based on the homogeneity of elite users, in which celebrities tend to follow celebrities and media tend to follow media. However, because the study focuses on the elite users, who are less than 1% of total users, there is a need to highlight the relations with users who follow these elite users. This current study analyzes the relations of exerting and receiving influence of opinion leaders, a similar concept to elite users. It also examines the patterns of influence between influential individuals and regular users through network analysis.

### 3. METHODOLOGY

This study addresses three research questions to investigate the communication patterns and network structures of the opinion leaders on Twitter during the 2011 Seoul mayoral elections.

First, this study will examine how the two candidates communicated with the public in the aspects of audience size and frequency using their Twitter usage patterns. Although this study focuses on Wonsoon Park, who actively used Twitter for campaigning purposes, it also refers to candidate Kyungwon Na for a basic comparison of the data.

RQ 1. Did the two candidates show different Twitter usage patterns during the elections?

This study finds the amount of influence through a quantification process of the data and also examines who these opinion leaders are and whether they are also influential offline. The reason we examine offline activities along with online activities is to check whether the selected opinion leaders continue their online activities accordingly offline with the same objectives. Through this, we expect to find significant differences between offline communities and virtual communities.

RQ 2. Did the opinion leaders actually have influence on the general public in the virtual community during the election period?

The last question is about the communication patterns of opinion leaders and the general public. This study examines the structure of the connections of influential individuals and the public, their conversation structures, and whether the communication was mutual. Because this study analyzes the structure of communication on Twitter, investigating how these network structures were formed, whether they are unidirectional or reciprocal, and what characteristics they have, is the core of this study. Thus, we form the next research question as shown below.

RQ 3. What kinds of network structures do the communication between the opinion leaders and the general public form?

In order to answer the addressed research questions, user tweet data was gathered using the Twitter API. Using this data, we measured the influence indicators and applied the HITS algorithm and examined the distribution of the 16 types of social network triads defined by Holland and Leinhardt [3] in order to analyze the structure of tweets between the opinion leaders and the public. The gathered data con-

sists of a total of 19,277 tweets. Among these tweets, there were tweets of regular people that were responded to by one of the 17 original users being studied (candidate Na, candidate Park, and Park’s 15 mentors) whose data was gathered. These tweets were also exposed inside the original tweets, resulting in tweet data from 8,547 user IDs.

To gather the data, we defined opinion leaders as the members of mayoral candidate Wonsoon Park’s Twitter mentors. The data (shown in Table 1) includes tweet information of the two mayoral candidates -Kyungwon Na and Wonsoon Park- in addition to 15 members of Park’s campaigners who actively shared their opinions on Twitter. One notable fact is that Park specifically created a Twitter election camp to communicate directly to the public. On the other hand, candidate Na used Twitter mostly to inform her election pledges to the public. Data was gathered one month prior to the election date, from September 26 to October 26.

	Kyungwon Na	Wonsoon Park	Mentors
Tweets	215	1,336	17,676
Users	312	656	7,579

Table 1: Basic information of the collected data

In order to analyze the network structure and communication patterns while also visualizing the network, the Fruchterman and Reingold algorithm was used [1]. While the user clusters and network structure was visualized with the Fruchterman and Reingold algorithm, the communication patterns between individual users in those clusters were analyzed using social network triads. These triads are frequently used to analyze network structure, and were used in this study to examine the most frequent communication patterns on Twitter during the election campaigns.

## 4. RESULTS

### 4.1 Relation Patterns between Opinion Leaders and the Public

Before analyzing the communication patterns of both candidates on Twitter, we investigated basic information on the candidates’ Twitter usages. Table 2 shows the information on candidate Na and Park’s number of followers, number of users they follow, tweets, lists, and join dates. Despite candidate Park joining Twitter two months later than candidate Na, he was a heavy user, having three times more followers and writing 12 times more tweets.

	Kyungwon Na	Wonsoon Park
Followers	57,425	189,623
Following	23,472	32,975
Tweets	860	9,845
Listed	2,867	7,197
Join Date	22nd July, 2009	22nd September, 2009

Table 2: Twitter usage data of the two candidates

In addition to examining the basic factors that related to the communication range of each candidate, the categorization of Twitter usage of the candidates was performed (shown in Table 2). There are two categories for ‘purpose of usage’ and two categories for ‘method of conversation’

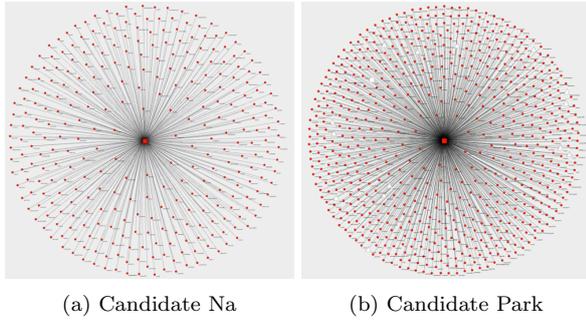


Figure 1: The communication range of candidate Na and candidate Park.

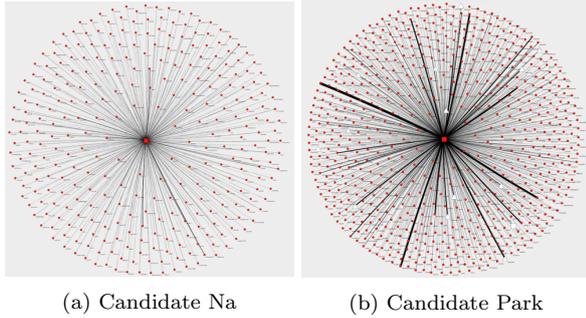


Figure 2: The communication frequency of candidate Na and candidate Park. Thicker edges indicate higher frequencies.

during the campaign period (September 26, 2011 ■ October 26, 2011). During the election campaign, official tweets from the candidates were required to insert a header in the tweet, stating that the tweet was for campaigning purposes. Thus, tweets that had no special headers or statement of purpose were categorized as regular tweets. For the conversation patterns, tweets that contained '@' were classified as Mentions and those that contained 'RT' were classified as Retweets. The conversation pattern shows the volume and frequency directly connected to the public. Both of the candidates used Twitter for the purpose of campaigning. However, Park had more conversations, particularly in the case of Retweets. Candidate Park usually used the method of retweeting to converse with Twitter users, at the same time enabling all of his followers to receive the interaction between Park and other ordinary users.

A basic analysis on the two most influential leaders was done regarding the audience size and frequency of communication. The audience size shows the volume of communication done between a candidate and the public, quantitatively. The frequency shows how often a candidate communicated with users from the public. To visualize the communication structure between the candidates and regular users, the Fruchterman and Reingold graph layout algorithm was performed on the social network graph. The nodes in the center of the two graphs represent the two candidates. Circle nodes represent the public and square nodes represent the mentors who are major opinion leaders on Twitter.

A quick comparison of Figure 1a and Figure 1b shows that the audience size of Park is significantly larger than that of Na. This result comes from several factors, such as the

volume of usage, the purpose of usage, and familiarity with Twitter. Figure 2a and Figure 2b show the frequencies of the number of conversations between the candidates and the public. Higher frequencies between two nodes are indicated with thicker edges.

Not only does Park have a larger audience than candidate Na, but also has a higher frequency of communication. One notable observation is that Park is connected to a large number of circle nodes, indicating that Park communicated with the general public directly, in addition to his mentors, during the campaign period. When looking at the Twitter conversations between the two candidates and the public in terms of structure, it was found that candidate Park communicated with a larger audience and a higher frequency than candidate Na.

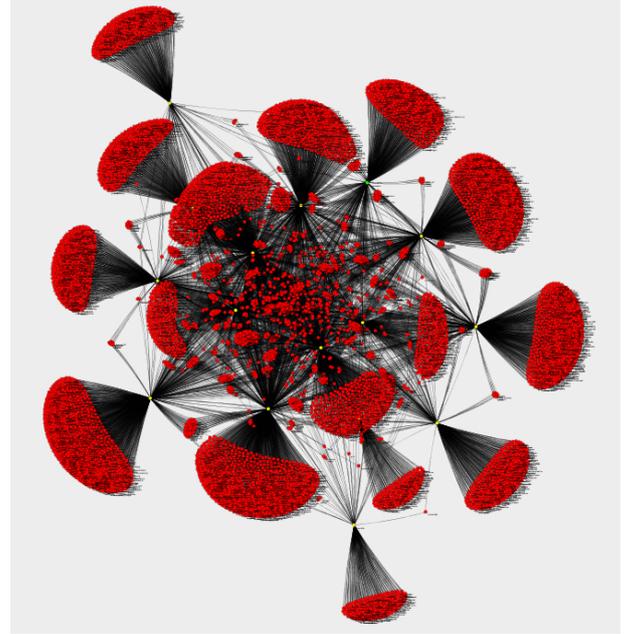


Figure 3: The structure of the communication network between Twitter users surrounding the central figures of the 2011 Seoul mayoral elections

We visualized the network patterns using the F&R algorithm on the collected data. While the aforementioned basic analysis on the data was about the two most influential candidates, this analysis is an analysis on the entire formation of the virtual community, which includes opinion leaders and those from the general public that they have connections with. Figure 3 shows the clusters of users who communicated with candidate Park on Twitter. Each circle node represents a twitter user. The square node represents mayoral candidate Park, and the triangle nodes represent the mentors of Park.

A notable finding is that the mentors (the triangle nodes) act as gateways: the triangle nodes gather and connect the clustered public users. This indicates that the mentors are the center of the network and also structurally connect the general users, acting as gateways to clusters of regular users.

## 4.2 Influence Indices of Twitter Opinion Leaders

HITS, or Hyperlink-Induced Topic Search, is a link analysis algorithm that was originally designed to rate web pages. When a user issues a search query, HITS first expands the list of relevant pages returned by a search engine and then produces two rankings of the expanded set of pages, authority ranking and hub ranking [6]. An authority is defined as a page with many in-links. The idea is that a webpage that has good or authoritative content will have many people that trust it and link to it. A hub is defined as a page with many out-links. Hub pages serve as an organizer of the information and links to many good authority pages. When users come to a hub page, they will find many useful links that take them to good content pages on the topic. The key idea of HITS is that a good hub points to many good authorities and a good authority is pointed to by many good hubs ([6]).

The first step of the HITS algorithm is to retrieve the set of results to the search query. In this study, the results of the ‘search query’ is explicitly given, i.e. the set of Twitter users that have been mentioned or whose tweets have been retweeted by the key members (who are also included in the set) of the mayoral elections. Authorities and hubs and their numerical values are defined in a mutual recursion. Simply put, a node’s, or a person’s, authority value is calculated to be the sum of the normalized hub values that point to that node, i.e. have a directed edge to that node. Likewise, a node’s hub value is computed to be the sum of the normalized authority values that point to that node. Thus, a node is assigned a high authority score when it is linked to by nodes that are considered hubs. Similarly, a node is assigned a high hub score when it links to nodes that are considered to be authorities.

Because the algorithm uses values that are calculated recursively, all nodes are given a hub score and authority score of 1 on the first iteration. Then, for each node in the graph, its authority and hub score is updated, according to the method stated above. The authority and hub scores are then normalized by dividing the authority scores by the sum of the squares of all authority scores and dividing the hub scores by the sum of the squares of all the hub scores (shown in Figure 4 and Figure 5). This process is repeated until the hub and authority scores converge to their final values.

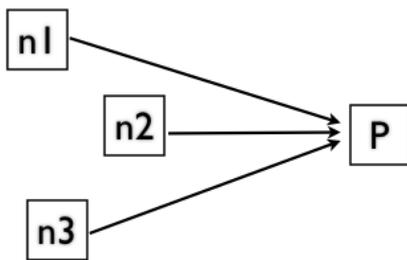


Figure 4: The Authority score for node P is the sum of the Hub scores of all nodes that point to node P.

This study finds the amount of influence of certain users through a quantification process by adopting the HITS algorithm to identify authorities and hubs in Twitter during the Seoul mayoral elections. The offline media exposures of these authorities and hubs are then examined and compared with their authority and hub values. Media exposure

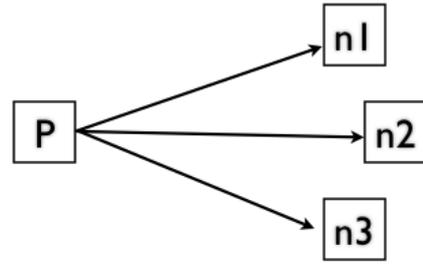


Figure 5: The Hub score for node P is the sum of the Authority scores of all nodes pointed to by P.

is defined as the number of newspaper articles from major newspapers that report on the opinion leaders during the Seoul mayoral elections. The newspaper articles were searched using the Korean Integrated Newspaper Database System (KINDS) homepage. In this study, the results of the HITS algorithm and frequency of media exposure (shown in Table 3) were compared in order to see whether opinion leaders, who are active on Twitter, also have influence on people offline by creating news value.

Table 3 shows the media exposure and authority and hub values of the mayoral candidate Park and Park’s mentors. The entries in bold are the top three highest-ranking items of each section. An interesting point is that the majority of opinion leaders, except candidate Park, were rarely exposed to offline media. The top five authorities, except Jeeyoung Gong who had published her book very recently, had little news value to offline news organizations.

Twitter user ‘Coreacom’ holds the highest rank of the authority playing the role of MeMedia on Twitter. However, he never had actually been exposed to offline media during the election period. User ‘Du0280,’ who is ranked the second in authority scores, had introduced himself as a Twitter news messenger. He also was rarely exposed to offline news organizations. User ‘patriamea,’ who is a professor at Seoul National University, is ranked the highest in hub scores. He is mentioned frequently in offline news organizations because he had created social issues for participating politically as a professor. User ‘welovehani,’ a newspaper reporter who was not mentioned in offline media, was second in the ranking in hub scores.

It can be seen that a person who does not necessarily receive much attention offline can have a significant news value and influence on the general public as an opinion leader in virtual communities, such as Twitter. ‘Coreacom’ had expressed himself as MeMedia, and ‘du0280’ had expressed himself as a Twitter news messenger. and ‘welovehani’ had introduced himself as a news megaphone for the powerless ordinary people (shown in Figure 6). They had made full use of Twitter as a new alternative media speaking for the people.

### 4.3 Transitivity in the Network of Communication Flow

The 16 types of social relations by Holland and Leinhardt [3] are usually adopted to analyze patterns of network structure [8]. Figure 7 shows the 16 different triads for directed graphs.

Figure 7 shows 16 types of triads by Holland and Lein-

Name	ID	Followers	Media Exposure	Authority	HUB
Sangmin Han	coreacom	201,020	0	<b>0.9567</b>	0.0018
Youngsuk Seo	du0280	17,675	4	<b>0.2063</b>	0.1277
Jeeyoung Gong	congjee	<b>217,625</b>	<b>172</b>	<b>0.1328</b>	0.0357
Sunggeun Moon	actormoon	100,101	25	0.1030	0.0728
Jaehyun Huh	welovehani	48408	0	0.0771	<b>0.2580</b>
Jaeyul Go	dogsul	134,965	2	0.0725	0.0177
Wonsoon Park	wonsoonpark	189,623	<b>1,958</b>	0.0251	0.1056
Yohjin Kim	yohjini	136,536	42	0.0248	0.0428
Jungkwon Jin	unheim	149,521	16	0.0240	0.0128
Guk Jo	patriamea	163,059	133	0.0180	<b>0.3757</b>
Junghee Lee	heenews	102,107	91	0.0160	<b>0.1398</b>
Hyeshin Jung	mindjj	45,136	10	0.0128	0.0130
Oisoo Lee	oisoo	<b>995,551</b>	39	0.0062	0.0231
Jedong Kim	keumkangkyung	<b>599,304</b>	31	0.0057	0.0346
Bongjoo Jung	BBK_Sniper	130,644	71	0.0056	0
Kyungmin Shin	mentshin	44,535	20	0.0002	0.0173

Table 3: Triad occurrences in the network of Twitter users of the 2011 Seoul mayoral elections



Figure 6: Twitter profiling page of ‘coreacom’ (1st in authority score), ‘du0280’ (2nd in authority score), and ‘welovehani’ (2nd in hub score)

hardt [3]. The coding refers to the numbers of mutual, asymmetric, and null dyads, with an identifying letter of Up, Down, Cyclical, Transitive. In the case of 021D, this triad has 0 mutual dyads, 2 asymmetric dyads, 1 null dyad, and in the Down orientation. Here, transitivity of a relation means that when there is a tie from  $i$  to  $j$ , and also from  $j$  to  $k$ , then there is also a tie from  $i$  to  $k$ .

Transitivity digraphs are important because it is a key structural property in social network data. The transitivity property of a social network, the most stable structure comprised of three nodes and three links, is an appropriate tool for analyzing the structure of the network or the patterns of communication within that network.

This study analyzed the network of the gathered tweet data and identified the distribution of occurrences of the above-mentioned 16 types of triads. Table 4 shows the results of this analysis. The six most frequent triads were Types 1, 2, 4, 5, 12, and 9 (shown in Figure 7). Type 1 has no mutual and asymmetric dyads. Type 2 has one asym-

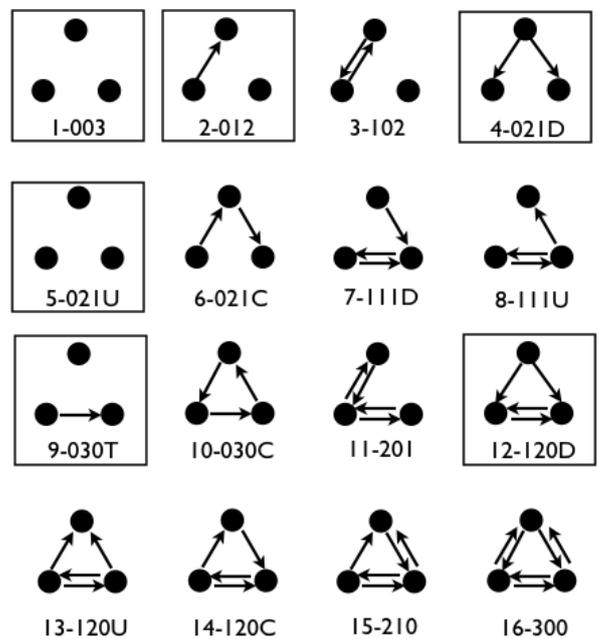


Figure 7: The 16 types of triads defined by Holland and Leinhardt, with the most frequent triads found in Twitter indicated with boxes.

metric and two null dyads. The most frequent two types of dyads do not have any specific directivity, which means the users are dispersed in the Twitter space. Type 4 has no mutual relation, having two asymmetric and one null dyad, with a downward relation. Type 5 has the same dyads with Type 4, but it has the upward direction. Type 12 has one mutual and two asymmetric dyads and a downward direction. Lastly, Type 9 is the only transitive structure among the six most frequent triads. It has three asymmetric dyads.

An interesting point to note is that out of the six most frequent triad types, only one has transitivity, which is contrary to the fact that Twitter is known as a space where communication is free. The network was mostly comprised of Type 1

Triad Type	Occurrence
1	136,054,019
2	88,485,118
3	353
4	4,814,803
5	412,769
6	2
7	35
8	4,536
9	22,123
10	22
11	124
12	86,149
13	43
14	13
15	106
16	1,263

Table 4: Triad occurrences in the network of Twitter users during the 2011 Seoul mayoral elections

and Type 2 triads, indicating that users on Twitter were disconnected and fragmented. As shown by the triads Type 4, 5, and 12, the up and down directionality means that opinions spread from or converge to a single node, rather than circulating through multiple nodes. This node, in which opinions spread from or converge to, is highly likely to represent an opinion leader. In other words, although Twitter’s network pattern shows a case of transitivity, such as the Type 9 triad, the majority are structurally fragmented or one-way conversations. Twitter is regarded as a public sphere under the new SNS environment. However, communication happens in the form of aggregation and propagation rather than in the form of sharing and circulating of various ideas in a society.

## 5. DISCUSSION

In this study, we examined what affects influential individuals had over the public. In addition, we examined the contribution these effects had on the formation and development of a virtual community. In particular, through the 2011 Seoul Mayoral Elections, we examined how opinion leaders communicated and what kind of network patterns were formed in Twitter’s virtual space.

Through a simple analysis of communication patterns, this study found that candidate Park had a significantly wider communication range (larger audience) and higher frequency compared to candidate Na. Park carried on his conversations with, not only the opinion leaders, but also with the general public, frequently. The opinion leaders had the roles of ‘gateways’, which concentrated and integrated the ideas in the Twitter space during the campaign period. Also, the HITS algorithm was performed to examine the authorities and hubs that play a major role in information propagation. The results show that a person who does not receive much attention offline can still have a significant news value and have influence on the general public as an opinion leader on Twitter. For the individual analysis of communication patterns using the 16 types of social network triads, it was found that the majority of the network patterns are fragmented or one-way conversations. Thus, although candidate Park tended to use Twitter actively as usual, the

results of analyzing the tweets and communication with the public during the mayoral elections show that the communication patterns of Park were mainly fragmented. Therefore, when excluding the activities of the mentors and just comparing the two candidates, there is no significant difference in the network patterns in communicating with the public. In addition, this signifies that the Twitter mentors play an important role as the aforementioned ‘gateways’ to the public for candidate Park.

In Twitter, we see that opinion leaders communicate with the public by reinforcing their opinions, rather than crystallizing them. In particular, there is a tendency for like-minded people to gather together and concentrate and integrate certain perspectives rather than circulating and sharing various opinions. Thus, while the influence of traditional opinion leaders was broad and reached many people, the influence of opinion leaders on Twitter seem to be more limited, only reaching out to clusters of similar users. This research is significant in that it attempts to reinterpret the quantitative analysis methods of social network services using social science as the framework. In particular, a multi-disciplinarian approach was taken through using network analysis methods, such as the HITS algorithm and social network triads, on traditional communication models.

Nevertheless, this paper has several limitations. First, the gathered data was limited to the tweets of the few, hand-picked opinion leaders during the election campaign period, instead of the entire tweets. Because the edges of the social network graph were created using communication patterns between users that used retweets or mentions, the graph may not be completely representative of the users’ actual communication patterns. In addition, this study would have been more significant if it were possible to measure the changes in the attitudes of users that resulted from the influences of opinion leaders during events from the real world. Also, a more in-depth study on the actual influence of virtual communities on the behavior in offline communities should be accompanied.

## 6. REFERENCES

- [1] T. M. J. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Softw. Pract. Exper.*, 21(11):1129–1164, Nov. 1991.
- [2] B. Heil and M. Piskorski. New twitter research: Men follow men and nobody tweets. *HBR Blog Network*, June 2009.
- [3] P. W. Holland and S. Leinhardt. Local structure in social networks. *Sociological Methodology*, 7(1976):1–45, 1976.
- [4] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web, WWW ’10*, pages 591–600, New York, NY, USA, 2010. ACM.
- [5] P. F. Lazarsfeld, B. Berelson, and H. Gaudet. *The People’s Choice. How the Voter Makes up his Mind in Presidential Campaign*. Columbia University Press, New York, 1944.
- [6] B. Liu. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Springer, 2007.
- [7] X. Song, Y. Chi, K. Hino, and B. Tseng. Identifying opinion leaders in the blogosphere. In *Proceedings of*

- the sixteenth ACM conference on Conference on information and knowledge management, CIKM '07*, pages 971–974, New York, NY, USA, 2007. ACM.
- [8] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.
- [9] D. J. Watts and P. S. Dodds. Influentials, Networks, and Public Opinion Formation. *Journal of Consumer Research*, 34(4):441–458, Dec. 2007.
- [10] J. Weng, E.-P. Lim, J. Jiang, and Q. He. Twitterrank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining, WSDM '10*, pages 261–270, New York, NY, USA, 2010. ACM.
- [11] S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts. Who says what to whom on twitter. In *Proceedings of the 20th international conference on World wide web, WWW '11*, pages 705–714, New York, NY, USA, 2011. ACM.