

# The Talkative, the Popular, and the Influential Korean Twitter Users

Hyun-joo Kong<sup>1</sup>, Hyun-woo Park<sup>1</sup> and Sang-ki Steve Han<sup>1</sup>,

<sup>1</sup> GSCT, KAIST, Yusung-gu, Gu-seong-dong, 373-1,  
305-701 Daejeon, Korea  
{hkong, shineall, stevehan}@kaist.ac.kr

**Abstract.** This study presents observation, analysis and discussion on various types of top Korean Twitter users. We focused on the use of communication conventions and functions for the clustering and conducted further analysis on them. We found that among the top 5% Korean Twitter users, there is no statistically significant relation between the number of tweets and followers meanwhile there is among the total 7,703 users. Distinctive characteristics among the three types of top users are discussed.

**Keywords:** microblogging, Twitter, use behavior, influence, popularity

## 1 Introduction

Twitter, a microblogging service, allows to its users a high degree of freedom in the service use by providing them with simple functions and unique restrictions in communication. In this paper, we observed the use of functions for communication among top Korean Twitter users to argue that the use patterns are related to certain propensities of users in the service use. We defined the top Korean users with various criteria and compared with each other. We collected data of 7,703 users and 1.3 million tweets (1,306,804 tweets) generated by them for 50 days (8/20/09-10/14/09).

## 2 Functions for communication in Twitter

**Length Restriction.** Twitter is a social network based microblogging service with a restriction of 140 maximum characters on a post, a tweet. This restriction on the length of each tweet makes a tweet very clear about its communicative purpose. Each tweet can serve only limited number of communicative purposes for instance, exchanging social conventions to show intimacy to one's followers and providing a new piece of information with comments is hard to be carried out. Messages should be expressed succinctly.

**Communication practices: Reply and Retweet.** Once a tweet is posted, it is displayed on the followers' timelines of the user. After receiving and reading a tweet, a follower may reply to the tweet. Motivations of replying to a tweet are various. A user may find a certain tweet especially interesting or he may spontaneously feel like to show some reaction to the user of the original tweet. Compared to retweet, which will be discussed later, Twitter automatically erases the original tweet. Therefore, it can be assumed that replies are more closely related to social relation with other users than to the tweet content itself.

A reply tweet typically starts with '@' followed by the ID of the user, who posted the original tweet. This form, '@+ID', is a convention for referring a user in Twitter. The syntax is used simply to mention other user's ID. Having one's user ID followed by '@' is important to a twitter user because it enables the ID relatively easily accessible to other users and possibly leading to new follower relations. [1] discussed the conversational aspects of reply tweets.

A follower can also retweet a tweet, which is called 'RT'. Different from a reply tweet, a RTed tweet can reserve its original tweet as well as the user ID. This convention leads it to serve various functions, for instance, showing sympathy, or acknowledgement on a value of a certain tweet rather than the user of the original tweet. [2] discusses various usages of RT.

**Nonreciprocal Relation.** Users in Twitter are tied to each other with follower and/or following relationships; inbound and outbound respectively. The relationships between users are not necessarily reciprocal. Following relationship enables a user to read his following users' tweets. In follower relationship, a user does not have control over his followers except 'block', which would be a social blow to the blocked user. Because relationship in Twitter is not necessarily reciprocal, Twitter has a sufficient condition to easily develop stars in its social networks.

These special characteristics cause one more difference between the actions of RT and reply. A RTed tweet is reached to the union of the followers of the two users; one who produced the original tweet and the other who RTed it. Instead, a replied tweet is delivered to the intersection of the followers of the two users, i.e., followers of both of the two users. Considering two factors, firstly the originality, whether a posted tweet is generated by the user himself or by other people and secondly, the openness, how far it can be delivered, we could categorize communication practices as follows.

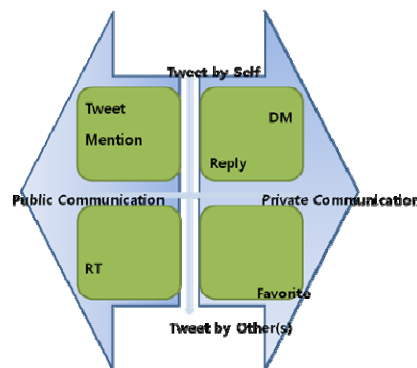


Figure 1. Categorization of communication practices by openness and originality

### 3 Data Collection

**Korean User.** [3] pointed out that Asian and European communities have a higher degree of correlation and reciprocity than communities in other continents and he attributed it to language barrier. Moreover, according to [4], almost half or a third of the top 25 ranks, who received the biggest number of replies in the world, were almost always occupied by Koreans. Considering that the Korean users occupy the 1.3% of the overall Twitter users [5], [6], this observed level of activity is extremely high. We decided to conduct analysis on the use behavior of top Korean twitter users.

**Data Collection.** We collected data of 7,703 users including 1.3 million tweets (1,306,804 tweets, specifically) generated by them for 50 days (8/20/09-14/10/09). To collect only Korean user data, we used the snowball sampling method [7]. We first identified 20 Korean users from a Korean twitter user website. From them, we expanded the number of subjects using their following relations filtering Koreans with location information and a map API provided by Yahoo.

**General Users.** General usage patterns of Korean Twitter users are as follow: 51.5% of the total tweets contain @, 16.5% of them URL, and 15.8% 'RT'. Fig. 2. compares Korean use patterns with those of international users discussed in [2]. Fig. 2. shows that the distinctive characteristics of Korean users are firstly the frequent use of '@' and 'RT' and secondly relatively small use of simple tweet, daily chat, which has no '@', URL, 'RT' etc. From those, we can say that Korean users are conversation-oriented Twitter users and they show strong tendency for interaction with other users.

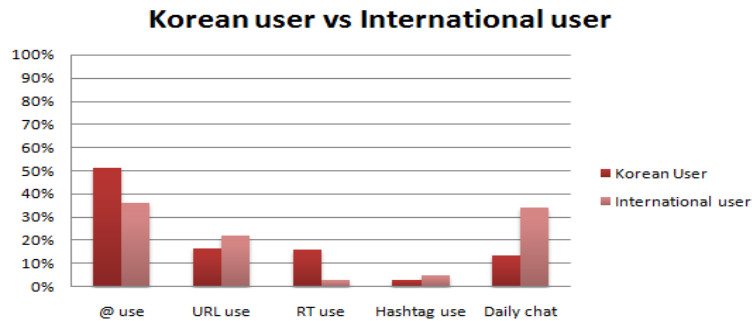


Figure 2. Comparison of Korean users' use patterns with those of international users

Usually the top 20% of user are reported to generate 80% of the total tweets. However in Korea, the top 14% of users account for 80% of the total tweets and the top 5%, the 57% of the total tweets.

**Top Users.** We identified top Korean users by three different criteria: the number of tweets, the number of followers and the level of influence. They represent active, popular, and influential user groups respectively. We defined an influencer as a user, whose tweets trigger further actions such as RTs and replies from other users. From this definition, the third group of influencers was identified by the each user's ratio of RTed tweets compared to the total RTed tweets and the replied tweets compared to

the total replied tweets. We used Ward clustering algorithm to identify groups. Resultantly, the top 9%, 9.5% and 5.03% of users were clustered as active, popular, and influential users respectively.

Table 1. Use behavior comparison among the three types of top Korean users

	Tweet	URL	@	RT	Hashtag
Active	57.21%	50%	66.98%	55.93%	43.81%
Popular	44.39%	26.44%	54.07%	50.81%	35.41%
Influential	52%	25%	65%	57%	36.70%

Table 2. Use behavior comparison among the three types of top Korean users

	Replied	RTed	RTed (w/ URL)	Rted (w/o URL)
Active	70.23%	67.67%	59.15%	40.55%
Popular	66.76%	72.81%	72.32%	38.53%
Influential	76.28%	78.12%	72.82%	41.49%

Active users are found active in the overall communication practices such as producing tweets, and using URL, @, RT and hashtags. In our sample, active and influential users are found to be conversation-oriented. Influential users produce more tweets and those with '@' than popular users. They also show selective uses of URL (25%) which lead to astonishing rate of being RTed (72.82%). We conducted cross-tabulation between the 'influential' and the 'popular' user groups, and the result showed that the majority of higher popularity (40 out of 57) fell in a lower level of influence.

Generally, it is pointed out that there is a strong positive relation between the number of tweets and followers. However we found that the relation does not exist among the top Korean users. In the top 9.5% users by the number of followers, there was no relation between the two variables (correlation coefficient -0.042). However, the relation between the number of tweets and followers remains positively strong (correlation coefficient 0.157) among the total Korean users. Fig. 3. shows the comparison between the influencer group and the total Korean users regarding the relation between the number of tweets and followers.

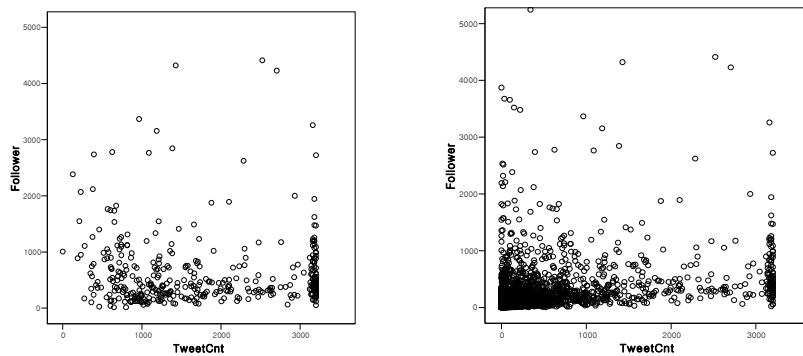


Figure 3. Correlations between followers and tweet counts in top 5% influential users (left) and in total sample (right).

## 4 Discussion

We examined the use practices of Korean Twitter users and focused further on the three types of top Korean Twitter users. Main findings are first, Korean users show direct interaction-oriented Twitter use. From this, we could infer that they are conversation-oriented users rather than information-sharing. Second, active users, who are identified by their number of tweets, found to be most active in overall communication practices such as the tweet generation, and the use of '@', 'RT', and hashtags. Third, URLs once introduced by influential users are highly likely spread over to other users. Finally, there is no relation between the high productivity of tweets and being influential or popular.

This study is a preliminary research to define and identify influencers in the microblogging service sphere especially among Korean users. In this study, influencers are defined by the quantity of responses from other users. In the future work, we will try to classify responses to achieve more detailed understanding on influencers. In this study, influencers among Korean Twitter users are found to show communication-oriented service use. From this, we will further examine a way to distinguish influencers who show information-sharing service use. On top of these, comparisons with users of other cultural backgrounds are expected to be an interesting research topic.

## 5 References

1. Honeycutt, C., Herring, S.: Beyond Microblogging: Conversation and Collaboration via Twitter. *Proc. HICSS-42*, IEEE Press (2009)
2. boyd, d., Golder, S., Lotan, G.: Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. *Proc. of HICSS-43*, IEEE Press (to be published)
3. Java, A., Song, X., Finn, T., Tseng, B.: Why we Twitter: Understanding microblogging usage and communities. *Proc. Joint 9th WEBKDD and 1st SNA-KDD workshop*, ACM Press (2007)
4. twitoaster. <http://twitoaster.com/>
5. Rankey. <http://www.rankey.com/>
6. ComScore. <http://www.techcrunch.com/2009/08/03/twitter-reaches-445-million-people-worldwide-in-june-comscore/>
7. Ahn, Y-Y., Han, S., Kwak, H., Moon, S., Jeong, H.: Analysis of Topological Characteristics of Huge Online Social Networking Services. *Proc. WWW*, ACM Press (2007)