8-2016

# Extending the network: The influence of offline friendship on Twitter network

Young Soo KIM
*Singapore Management University*, yskim@smu.edu.sg

Kyungsub Stephen CHOI

FELICIA NATALI
*Singapore Management University*, felician.2013@phdis.smu.edu.sg

# Extending the Network: the Influence of Offline Friendship on Twitter Network

*Full papers*

**Youngsoo Kim**
School of Information Systems
Singapore Management University
Singapore 178902
yskim@smu.edu.sg

**Kyungsub Stephen Choi**
School of Management
Rhode Island College
Providence, RI, USA 02908
kchoi@ric.edu

**Felicia Natali**
School of Information Systems
Singapore Management University
Singapore 178902
felician.2013@smu.edu.sg

## Abstract

Twitter effectively provides a communication platform that allows users to strengthen or augment relationships with their close ones. It is a common scene to see groups of people in continuing group communication from offline to online using tools such as Twitter. Furthermore, it is also possible to meet new friends via online encounter. Under these circumstances, there would be friends who are both offline and online, or only in online. Many of earlier Twitter network studies focused on the network effects' direction going from online to offline network. This paper explores the opposite direction, going from offline to online network. It investigates the peer's friendly relationship intricacies that emerge when a known friendly offline relationship influences its subsequently established Twitter online relationship. An empirical data set of 2,193 pairs of Twitter user accounts were examined.

**Keywords**
Twitter, Offline relationship, networking behavior, social network, and online friendship

## Introduction

As one of the leading social media, Twitter is a well-established world-wide social medium. According to twitter.com webpage, there are 320 million monthly active users, 80% of them are on mobile platform, 79% of them are outside the United States, and available in more than 35 different languages. Started out as a communication service for small group setting (Miller 2008), Twitter is general known as a tool for status updates and social utility. However, given its versatility, it appeals for more than its nominal function (Lapowsky 2013; Crawford 2009; Kwak et al. 2010).

Transcending the location and time barriers, Twitter effectively provides a communication platform that allows users to strengthen or augment their relationships with their close ones. Furthermore, Twitter also provides a ground for users to meet new users where they would not have a chance to meet because of the temporal and proximal restrictions. These Twitter's goodness well buttresses a person's offline friendly network. Not only Twitter can augment a person's existing offline network, but it can also extend a person's network by adding new friends into the existing network. Under these circumstances, there would be friends who are friends both offline and online, or friends in only online.

This paper investigates the peer's friendly relationship intricacies that emerge when a known offline friendly relationship influences its subsequently established Twitter online friendly relationship network. This is achieved thru a quantitative analysis on an empirical Twitter data set. Many of the published earlier Twitter quantitative analysis studies have examined the *opposite* direction: how online network impacts offline network. Few of those studies are: derivation of one's private offline information (political affiliation) from online social network structure (Heatherly et al, 2013), investigation of how online friendship network structure reveals offline high-risk sexual behavior (Dai et al, 2012), creation of how an algorithm to predict the offline friends based on the Twitter connections (Xie et al, 2012), and formulation of a new network measure called dispersion – the extent to which two people's mutual friends are not well-connected – to predict the family members on Facebook (Backstrom and Kleinberg, 2014).

Albeit a person may encounter strangers online, the fact about Twitter is that, as a status update utility, people generally move from offline to online in regards to network formation. The contribution of this study lies at understanding how this offline friendly network migrates and influences the establishment of Twitter network. For this, this study delves into (a) network structure, (b) content analysis and (c) interaction analysis using a regression model.

## Research Framework

There have been a number of diverse inquiries and studies about offline and online friendships using social media. Some of these studies would be examined to an extent that is relevant to this study. The relevant areas of the study we have looked into are: (1) Twitter network, (2) comparison between the offline and online friendship, (3) Homophily between friends, and (4) the examination of the influence of offline friendship on online networking behavior.

For the research question, we aim to the following question: "Given a pair of Twitter users, if two users know each other offline (offline friendship edge), do they show different networking behavior on the Twitter network compared to other pairs of users who do not know each other offline (online friendship edge)?" Specifically, we examine the influence of offline friendship on three specific dimensions of Twitter networking behaviors: (1) *network structure*, (2) *content similarity*, and (3) *interactions* on Twitter.

First, the definition of friendship in this study includes acquaintanceship and so we define offline friendship as a friendship between two users who know each other in the offline world. On the other hand, we define online friends as two users who are connected in Twitter (follower and/or followee) without any offline meeting. Here, when user *i* follows user *j*, user *i* is called user *j*'s follower, and user *j* is called user *i*'s followee. Although a Twitter user does not have a choice over his followers, the user may still get acquainted with his followers through replies or likes that his followers generate for him. In summary, users *i* and *j* are online friends in either of the following situations: user *i* follows user *j*, user *j* follows user *i*, or user *i* and user *j* follow each other.

Second, the scope of our analysis for the network structure is a Twitter ego network. A Twitter ego network consists of an ego user and all who are directly connected to the user, called *alters*. Figure 1 below illustrates a typical ego network of a Twitter user (user 1). We draw a link from user *i* to user *j* if user *i* follows user *j*. The connections are mostly between the ego user and his alters (followers or followees), but they also include all the connections among the ego's alters (e.g. between 3 and 6). A follow link goes against the information flow. Thus, if user *i* follows user *j*, information (tweet content) flows from user *j* to user *i*. In Figure 1, user 1 follows user 5. Information can flow from user 5 to user 1 on the Twitter network.

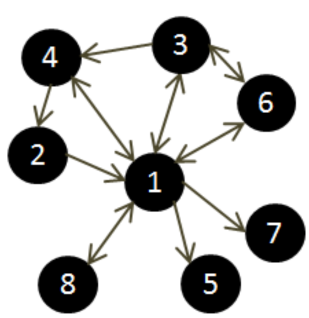


**Figure 1. Twitter Ego Network**

Various network measures have been developed to describe a network structure. We selectively choose and revise some of them to our research setting of Twitter to (1) understand the fundamental principles of social network formation (e.g., reciprocity and the number of mutual online friends) and (2) explore communication and information distribution on Twitter (e.g., network coverage ratio, information flow efficiency, and edge betweenness centrality). Additionally, we analyze Twitter content similarity and the interaction between users. Table 1 shows operational definitions of all the measures in this study and summarizes our findings.

| Category | Variable | Operational Definitions | Twitter Networking Behaviors Observed |
|---|---|---|---|
| | $j$ | ego user index | |
| | $i$ | alter index | |
| | $RelationType_{ij}$ | the relation type between user $i$ and user $j$ (offline or online) | |
| Network Structure | $Reciprocity_{ij}$ | whether the link from user $i$ to user $j$ is reciprocated | responding to online gesture of friendship |
| | $FollowerOverlap_{ij}$ $FolloweeOverlap_{ij}$ | followers/followees overlap between user $i$ and user $j$ | forming online friendships with whom one shares mutual friends |
| | $NCR_{ij}^{D}$ and $NCR_{ij}^{G}$ | user $i$'s network coverage ratio in distributing or gathering information in user $j$ ego's network | independence in distributing/gathering information |
| | $IFE_{ij}^{D}$ and $IFE_{ij}^{G}$ | user $i$'s efficiency in distributing or gathering information in user $j$'s ego network | independence in distributing/gathering information efficiently |
| | $EBC_{ij}^{D}$ and $EBC_{ij}^{G}$ | edge betweenness centrality of information distribution or gathering link between user $i$ and user $j$ in user $j$'s ego network | facilitating communication by: (a) posting tweets interesting to a friend, or (b) paying attention to a friend's tweets |
| Twitter Content Similarity | $Similarity_{ij}$ | content similarity between user $i$ and user $j$ | posting tweets similar to a friend's |
| Interactions on Twitter | $Favorite_{ij}$ | the number of user $j$'s tweets that user $i$ likes | liking a friend's tweets |
| | $Retweet_{ij}$ | the number of user $j$'s tweets that user $i$ retweets | retweeting a friend's tweets |
| | $Reply_{ij}$ | the number of user $i$ 's replies to user $j$'s tweets | replying a friend's tweets |
| | $Mention_{ij}$ | the number of user $j$'s names in user $i$'s tweets | mentioning a friend in tweets |

*D in superscript represents the Distribution flow of tweet.
*G in superscript represents the Gathering flow of tweet.

**Table 1. Variables, Operational Definitions and Twitter Networking Behaviors**

## *Network Structure*

### Reciprocity

Reciprocity is one of the fundamental principles of social network formation (Schaefer et al. 2010). In Twitter, reciprocity means following each other, a two-way friendship. When two friends follow each other, then each will receive the other's tweets and updates. Following a user on Twitter is a gesture of friendship and so reciprocity measures how close and intimate two users are. Granovetter (1973) has proposed reciprocity as one of the indicators of the strength of tie. Therefore, by exploring reciprocity, we can investigate the influence of offline friendship on the strength of online relationship on the Twitter network. The measure of strength tie is not only confined to reciprocity as there are many other indicators. Even reciprocity itself has several types (e.g., reciprocity in replying, mentioning). However, in this study, we focus only on structural reciprocity. A reciprocity between users $i$ and $j$ is measured by the following formula:

$$Reciprocity_{ij} = \begin{cases} 1 \text{ if users } i \text{ and } j \text{ follow each other} \\ 0 \text{ otherwise} \end{cases}$$

### Overlap of Followers and Followees

Another variable that is closely related to one of the fundamental principles of social network formation is the follower overlap and the followee overlap. As we described, the term *follower* is referring to other users who follow a user and the term *followee* is referring other users who are followed by the user. People have a tendency to form friendships with those whom they share multiple mutual friends. It is common that a group of friends form this multiple friendship among themselves. Therefore, in a social network, it is a highly probable that same individuals are followers and followees of two different users. The focus is the *overlap* of these followers and followees. This measures the extent to which two friends have the overlap of same followers and followees in their ego networks. This concept is called triadic closure, one of the fundamental principles of social network formation (Schaefer et al. 2010). By examining the overlap of followers and followees across users, we aim to investigate the influence of offline friendship on the configuration of online friendship network on Twitter. Since overlapping friends is an indicator of tie strength (Granovetter, 1973), we may also investigate tie strength by exploring the overlap of followers and followees. We measure *follower* overlap and *followee* overlap between user $i$ and user $j$ by the following formula (Easly and Kleinberg, 2010):

$$\boldsymbol{Follower}Overlap_{ij} = \frac{\text{The number of common followers of } i \text{ and } j}{\text{The number of unique followers of } i \text{ or } j}$$

$$\boldsymbol{Followee}Overlap_{ij} = \frac{\text{The number of common followees of } i \text{ and } j}{\text{The number of unique followees of } i \text{ or } j}$$

### Network Coverage Ratio

Network coverage ratio measures how *independent* a user is in distributing or gathering information to and from the alters. The alters are users to whom an ego user is directly connected by either following the user or being followed by the user. The independency is referring to whether a user can communicate directly to another user or indirectly through other users' connections. It is assumed that the more independent user $i$ is on user $j$ in distributing or gathering information to or from user $j$'s alters, the more likely user $i$ is in distributing or gathering information to or from user $j$'s friends because the information coming to or from user $j$'s friends is not subject to the influence of user $j$.

As information flows to the followers, in estimating how independent a user is in distributing information in his friend's ego network, we consider the number of reachable followers the user can reach. The formula calculates how independent user $i$ is from user $j$ in *distributing* information in user $j$'s ego network. When user $i$ is completely independent from user $j$ the value of $NCR_{ij}^{D}$ is 1. On the other hand, if user $i$ is completely dependent on user $j$, the value is 0.

$$NCR_{ij}^{D} = \frac{\text{The number of users in user } j\text{'s ego network who receive a user } i\text{'s tweet, when a tweet cannot flow through user } j, \text{ excluding user } j.}{\text{The number of users in user } j\text{'s ego network who receive a user } i\text{'s tweet, excluding user } j.}$$

## Information Flow Efficiency

Information flow efficiency measures how efficiently a user is propagating and gathering a tweet (information) in his friend's ego network. Similar to Network Coverage Ratio, Information Flow Efficiency assumes a hypothetical information distribution and not a real one. To remove the information benefit from ego-alter connections in quantifying the importance of an alter in the eyes of other alters, we consider the independence of an alter from an ego user in distributing or gathering information. With this measure, we attempt to evaluate the influence of offline friendship on efficient information propagation and gathering on Twitter. This measure is based on the notion of *closeness centrality* that was developed to calibrate the efficiency of information spread on network. A high closeness centrality translates to the minimum amount of time in spreading news (Bavelas 1950). The formula of closeness centrality uses the harmonic shortest distance measurement (Newman 2010).[1]

$$Closeness\ Centrality_i = \frac{1}{n-1} \sum_{k \neq i, k \in K} \frac{1}{Shortest\ Distance_{ik}}$$

$Closeness\ Centrality_i$ is the closeness centrality measure of user $i$, and $Shortest\ Distance_{ik}$ is the shortest distance (links) between user $i$ to user $k$ where $k \in K$ and $K$ can be defined as a set of any group of users. In Information Flow Efficiency formula, $K$ is user $j$'s alters. Based on the formula above, the formula for information flow efficiency in *distributing* information is provided below:

$$IFE_{ij}^{D} = \frac{Closeness\ Centrality_i \text{ in distributing tweets to user } j\text{'s alters in user } j\text{'s ego network, when tweets cannot flow through user } j}{Closeness\ Centrality_i \text{ in distributing tweets to user } j\text{'s alters in user } j\text{'s ego network,}}$$

The formula above calculates how independent user $i$ is from user $j$ in distributing information efficiently in user $j$'s ego network. When user $i$ is completely independent from user $j$, the value of $IFE_{ij}^{D}$ is 1. On the other hand, if user $i$ is completely dependent on user $j$, the value is 0.

## Edge Betweenness Centrality

Edge betweenness centrality measures how important a link between two users is in facilitating communications among peers in a network. Technically, edge betweenness centrality measures the extent to which a link in a communication network falls on the shortest path between pairs of other points (Girvan and Newman, 2002). In Twitter, a communication link is represented by a *following* link. Since a *following* link represents one-to-many instead of one-to-one conversation, in order for communication to flow between two users, a user either has to pay attention to his friend's tweets, or his friend has to post tweets that are of interest to the user. In this study, we are utilizing edge betweeness centrality to measure the influence of offline friendship on facilitating communication on Twitter through a follow link.

The formula for $EBC_{ij}^{D}$, the edge betweenness centrality of the communication link from $i$ to $j$ ($i$ Distributes tweet to $j$, user $j$ follows user $i$) is:

---

[1] In our measure, we consider 1-hop follow network for each user so that all users should at least have one-way follow link with the target user. This type of network is called an ego network. Note that a 2-hop Twitter follow network can easily have tens of millions of user nodes, or even more if large media/celebrities nodes are involved. Experiments from (Kwak et al. 2010) show that the average path length between any two users in Twitter is only 4.12, corroborating our setting of an ego network a small value.

$$EBC_{ij}^{D} = \frac{\sum_{s,t \in G_j} \frac{\sigma_{st}(CommLink_{ij})}{\sigma_{st}}}{n \times (n-1)}$$

Mathematically, $\sigma_{st}$ measures the number of shortest paths from user $s$ to user $t$. $\sigma_{st}(CommLink_{ij})$ is the number of shortest paths from user $s$ to user $t$ including the link from user $i$ to $j$. $G_j$ is the ego network of user $j$, and $n$ is the number of users in $G_j$.

### Content Analysis

Using the LDA approach, we extract topics out of tweets. We then measure the content similarity of topics extracted from tweets between two users to see if there is an influence of offline friendship on tweeting behavior, specifically how likely is it that a user posts tweets that are similar to his friend's. Using Latent Dirichlet Allocation(LDA) (Blei et al. 2003), we generate a topic distribution of a user's tweets. We calculate the similarity of tweets between two users by the following formula:

$$ContentSimilarity_{ij} = \frac{\exp\left(-D_{KL}(i|j)\right) + \exp(-D_{KL}(j|i))}{2}$$

$D_{KL}(i|j)$ is the Kullback-Leibler divergence (Kullback and Leibler 1951) of the topic distribution from user $i$'s tweets to user $j$'s tweets. $D_{KL}(j|i)$ is the Kullback-Leibler divergence of the topic distribution from user $j$'s tweets to user $i$'s tweets. Given $Z$ as a collection of topics,

$$D_{KL}(i|j) = \sum_{z \in Z} p(z|i) \times \log \frac{p(z|i)}{p(z|j)}$$

When the divergence value is divided by $\log 2$, it is converted to bits. In our experiment, we are using bits as the default unit of measurement.

### Social Interactions on Twitter

Myers et al. (2014) report that the Twitter follow graph exhibits structural characteristics of both an information network and a social network, showing that Twitter starts off more like an information network, but evolves to behave more like a social network. Some papers still refer to Twitter as a social network. Some tweets are known to be phatics in nature (Makice, 2009), promoting social interaction and the use of Twitter as a social network. As a social network, people use Twitter mainly to maintain connection with existing friends and meet new friends.

Twitter, as an online social network provides a platform for users to interact with one another. There are four types of interaction in Twitter: *favorite*, *retweet*, *reply*, and *mention*. 'Favorite' is used when a friend likes a tweet of another friend. 'Retweet' is reposting a friend's tweet. 'Reply' is responding back to a friend's tweet. 'Mention' is mentioning a friend in tweets. We develop the following measures to capture various social interactions in Twitter that are externalized through Twitter features such as *favorite, retweet, reply*, and *mention*.

$Favorite_{ij}$ = the number of user $j$'s tweets that user $i$ likes

$Retweet_{ij}$ = the number of user $j$'s tweets that user $i$ retweets

$Reply_{ij}$ = the number of user $i$'s replies to user $j$'s tweets

$Mention_{ij}$ = the number of user $j$'s name in user $i$'s tweets

### Regression Model

A regression equation was developed in order to run a number of possible outcome variables that would be affected by whether the two users were offline friends or not.

$$Outcome_{ij} = \beta_0 + \beta_1 RelationshipType_{ij} + \zeta_j + \mu_{ij}$$

The data structure for all the models was a cross-sectional data. On the outcome variables, the variables were pertained to (a) ego network properties, (b) tweets, and (c) interactions via Twitter. The equation was meant to explain the changes in $Outcome_{ij}$ per relationship type: online or offline. A fixed effect model ($\zeta_j$) are employed to control for user $j$'s network structure heterogeneity. The error component, $\mu_{ij}$ is an idiosyncratic error term and it varies across $i$ and $j$. If $\beta_1$ is positive and statistically significant, then it indicates that the offline relationship increases the strength of the measures.

## Data Collection

For the empirical data set, we contacted a considerable number of Twitter users and among them 98 users have consented and approved our data collection in their Twitter networks. Under their written agreements, we thoroughly examined each user account being an ego network and extracted information of (1) their *alters* and (2) content in Tweets that they and their alters have sent. To identify the offline friendship in their Twitter networks, we asked them whether they knew each Twitter user in their Twitter networks in real life. Additionally, we have excluded those types of users ineligible as a target user – e.g. spam users, dormant users, celebrities, politicians, etc. At the end, 2,193 pairs of Twitter online friends were identified and consequently included for the data collection. Among the 2,193 friends, 873 were also offline friends.

## Results and Discussion

Table 2 shows some descriptive statistics and correlation matrix for the variables that were used in the regression model. The correlation between the variables show whether there are any interesting relationships between the variables. The baseline correlations provide initial support for the general expectation, indicating that the users with offline relationship are positively associated with this study's measures except for $Favorite_{ij}$. However, the correlations cannot fully guarantee the causal relationship to the user's heterogeneity. Therefore, the regression model was administered with individual user fixed effect. Table 2 presents the regression results.

| $Outcome_{ij}$ | $\beta_1$ | Standard Error | # of Observations | Within $R^2$ |
|---|---|---|---|---|
| 1. $Reciprocity_{ij}$ | 0.35381*** | 0.01877 | 2193 | 0.1414 |
| 2. $FollowerOverlap_{ij}$ | 0.02211*** | 0.00144 | 2193 | 0.0987 |
| 3. $FolloweeOverlap_{ij}$ | 0.02129*** | 0.00174 | 2193 | 0.0652 |
| 4. $NCR_{ij}^{D}$ | 0.18919*** | 0.01496 | 2193 | 0.1663 |
| 5. $NCR_{ij}^{G}$ | 0.11289*** | 0.01514 | 2193 | 0.1127 |
| 6. $IFE_{ij}^{D}$ | 0.16002*** | 0.01274 | 2193 | 0.2298 |
| 7. $IFE_{ij}^{G}$ | 0.06728*** | 0.01344 | 2193 | 0.1396 |
| 8. $EBC_{ij}^{D}$ | 0.00083*** | 0.00010 | 2193 | 0.0379 |
| 9. $EBC_{ij}^{G}$ | 0.00070*** | 0.00018 | 2193 | 0.0098 |
| 10. $Similarity_{ij}$ | 0.11024*** | 0.008131 | 2193 | 0.0785 |
| 11. $Favorite_{ij}$ | 0.00704 | 0.00561 | 2193 | 0.001 |
| 12. $Retweet_{ij}$ | 0.14519* | 0.04817 | 2193 | 0.0048 |
| 13. $Reply_{ij}$ | 1.20393*** | 0.26508 | 2193 | 0.0102 |
| 14. $Mention_{ij}$ | 0.38023*** | 0.09533 | 2193 | 0.0075 |

*significant at $p$<0.1  **significant at $p$<0.01  ***significant at $p$<0.001

**Table 2. Regression Results**

The connecting mechanism on Twitter (following and being followed) does not provide a selective assessment on a friend, unlike on Facebook where a friend request needs to be approved selectively, unless the Twitter account is set to private in which case, a follower request needs to be approved. Reciprocity refers to returning a following request back to the sender. When a user follows a friend, the user also tunes into the friend's tweets. In Twitter, the gesture of friendship is represented by following a user. Reciprocity entails a positive or grateful response to the gestures of friendship by following a follower. In general, reciprocity is considered to be one of the building blocks of a social medium (Schaefer et al. 2010). The positive and significant coefficient of $RelationType_{ij}$ ($\beta_1$=0.354, $p$<0.001) for $Reciprocity_{ij}$ shows that offline friendship increases a user's likelihood to respond to another user's gesture of friendship online.

## *Overlap of Followers and Followees*

Our regression results show that two users with offline friendship are likely to share more friends in Twitter ($\beta_1$=0.022, $p$<0.001 for $FollowerOverlap_{ij}$ and $\beta_1$=0.021, $p$<0.001 for $FolloweeOverlap_{ij}$). Because strong ties have a higher neighbor overlap (Granovetter 1973). These findings corroborate the conclusion from the result on reciprocity – offline friends are stronger ties compared to online friends.

## *The Role of Offline Friendship in Information Flow on Twitter*

From the regression results for $NCR_{ij}^{D}$ and $NCR_{ij}^{G}$, the coefficients of $RelationType_{ij}$ show that offline friendships increase a user's independence in distributing or gathering information in his friend's ego network. Even a simple information such as price and store name can influence a buyer's product perception (Dodds, 1991). In reality, the information propagated on Twitter can be one's positive or negative opinion beyond objective knowledge. As such, our results indicate that a user is likely to extensively influences friends connected to an offline friend and acquire diverse information from friends of an offline friend.

In information flow efficiency, we consider only only the number of reachable users but also the distance of the reachable users. The positive and significant coefficients of $IFE_{ij}^{D}$ and $IFE_{ij}^{G}$ show that a user can more likely and quickly distribute or gather information in a offline friend's ego network. The results show an incentive for a user to establish online connection with more offline friends. As a user connect with more offline friends, a user may effectively and efficiently leverage Twitter network in distributing and gathering information.

## *Offline Homophily versus Online Homophily in Tweet Content*

When $Similarity_{ij}$ is regressed on $RelationType_{ij}$, we observe a significantly higher similarity level between offline friends ($\beta_1$=0.26, $p$<0.001). In offline case, the self-categorization theory tells that a person identifies with other people who are similar in the attributes that the person possesses (Turner and Oakes, 1986; 1989). Those attributes can be age, gender, education, social class, or occupation. Through one or more of these attributes, homophily is established (Carley, 1991). Sequentially, a tie develops and it gets stronger through higher frequency in social encounters. Hence, offline friends are likely to develop into strong ties. In online case, the self-categorization theory cannot be applied because a person cannot physically see the others in identifying the "similarities" using the above attributes (Katz et al, 2004). One exception is when people decide to reveal their offline identities in online (Hargittai 2007), but this is not a common behavior. Even when they do, identities can be distorted to display only what they want others to see.

In the online world, people mostly encounter the "similarities" through online cognition. An online cognition is an experience where a person "*cognitively connect*" to another person in an online context, conversation, or interaction. A comment may follows as "I don't know, but I just clicked with that guy, we chatted over hours and we felt like we knew each other for many years before." This happens randomly as it is unscheduled, unplanned, and unexpected. The feeling of comraderies and a sense of togetherness may take place. But on an average rate, this experience is rare to come by and a chance of this happening is minimal. In many cases, online encounters will remain in that one time session only.

## *Stronger Interactions with Offline Friends*

For the Twitter online group interactions, there were four measures: (1) '*favorite*' (now is called '*Like*' by Twitter.com), (2) '*retweet,*' (3) '*reply,*' and (4) '*mention.*' The '*favorite*' and '*retweet*' analysis results were found to be not so significant ($\beta_1$=0.00704, $p$>0.1 for $Favorite_{ij}$; $\beta_1$=0.14519, $p$<0.1 for $Retweet_{ij}$) whereas the '*reply*' and '*mention*' analysis results were significant ($\beta_1$=1.20393, $p$<0.001 for $Reply_{ij}$; $\beta_1$=0.38023, $p$<0.001 for $Mention_{ij}$). With these results, the question is why there were no significant results with

*Favorite*$_{ij}$ and *Retweet*$_{ij}$ when the group was composed with people who are both offline and online friends, and when offline friends also showed significantly higher *Similarity*$_{ij}$.

## Conclusion

Using an empirical data set of 98 proprietary Twitter accounts and 2193 user pairing, this study quantitatively analyzed the network structure, content analysis, and interactions of Twitter networking behaviors. In the network structure, these areas were investigated: reciprocity, network coverage ratio, information flow efficiency, and edge betweennes centrality. For the interactions, favorite, retweet, reply, and mention functions were examined. Through a regression model, each focused area or relationship were assessed. The regression results have revealed that except for *Favorite*$_{ij}$, the users with offline relationship are positively associated with all of the focused areas and relationships. Additionally, the study's results add to the theories of social network with the significances of offline friendship.

## Acknowledgement

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centres in Singapore Funding Initiative.

## REFERENCES

Backstrom, L. and Kleinberg, J. 2014. "Romantic partnerships and the dispersion of social ties: a network analysis of relationship status on facebook," in *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing,* Baltimore, MD, pp. 831-841

Bavelas, A. 1950. "Communication Patterns in Task-oriented Groups," *Journal of Acoustical Society of America* (22), pp. 725-730

Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. "Latent dirichlet allocation," The Journal of machine learning research (3), pp. 993-1022.

Carley, K. M. 1991. "A theory of group stability," *American Sociological Review* (56), pp. 331- 354.

Crawford, K. 2009. "These Foolish Things: On Intimacy and Insignificance in Mobile Media." in *Mobile Technologies: From Telecommunications to Media*, G. Goggin and L. Hjorth (eds.), New York: Routledge, pp. 252–65.

Dai, L., Luo, J. D., Fu, X., and Li, Z. 2012. "Predicting offline behaviors from online features: an ego-centric dynamical network approach," In *Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research,* Beijing, China, pp. 17-24.

Dodds, W. B. 1991. "In search of value: how price and store name information influence buyers' product perceptions," *Journal of Consumer Marketing* (8:2), pp. 15-24.

Easly, D., and Kleinberg, J. 2010. Networks, Crowds, and Markets. Reasoning about a Highly Connected World, Cambridge, UK, Cambridge University Press, 2010.

Girvan, M., and Newman, M. E. 2002. "Community structure in social and biological networks," in *Proceedings of the national academy of sciences*, (*99:*12), pp. 7821-7826.

Granovetter, M. S. 1973. "The strength of weak ties," *American journal of sociology* (78), pp. 1360-1380.

Hargittai, E. 2007. "Whose space? Differences among users and non-users of social network sites," *Journal of Computer-Mediated Communication*, (13:1), pp. 276-297.

Heatherly, R., Kantarcioglu, M., and Thuraisingham, B. 2013. "Preventing private information inference attacks on social networks," *Knowledge and Data Engineering, IEEE Transactions on* (*25:*8), pp. 1849-1862.

Hofer, M. and Aubert, V. 2013. "Perceived bridging and bonding social capital on Twitter: Differentiating between followers and followees," *Computers in Human Behavior* (*29:*6), pp. 2134-2142.

Katz, N., Lazer, D., Arrow, H., and Contractor, N. 2004. "Network theory and small groups," *Small group research*, (35:3), pp. 307-332.

Kullback, S., and Leibler, R. A. 1951. "On information and sufficiency," *The annals of mathematical statistics* (*22:*1), pp. 79-86.

Kwak, H., Lee, C., Park, H. and Moon, S., 2010. "What is twitter, a social network or a news media," in *International World Wide Web Conference*, Raleigh, NC. pp. 591-600

Lapowsky, I. 2013. "Ev Williams on Twitter's Early Years," Inc., Retrieved January 21, 2016, from http://www.inc.com/issie-lapowsky/ev-williams-twitter-early-years.html?cid=em01011week40day04b

Makice, K. 2009. "Phatics and the design of community," in *the Computer Human Interaction Conference Extended Abstracts on Human Factors in Computing Systems,* Boston, MA, pp. 3133-3136.

Miller, V. 2008. "New media, networking and phatic culture convergence," *The International Journal of Research into New Media Technologies*, (14:4), pp. 387-400.

Myers, S. A., Sharma, A., Gupta, P., and Lin, J. 2014. "Information network or social network? The structure of the twitter follow graph," in *Proceedings of the companion publication of the 23rd international conference on World Wide Web companion,* Seoul, Korea, pp. 493-498.

Newman, M. 2010. *Networks: an introduction*. 1st ed. Oxford, UK, Oxford University Press

Schaefer, D. R., Light, J. M., Fabes, R. A., Hanish, L. D., and Martin, C. L. 2010. "Fundamental principles of network formation among preschool children," *Social Networks*, (*32:*1), pp. 61-71.

Turner, J. C., and Oakes, P. J. 1986. "The significance of the social identity concept for social psychology with reference to individualism, interactionism, and social influence," *British Journal of Social Psychology* (25), pp. 237-252.

Turner, J. C., and Oakes, P. J. 1989. "Self-categorization theory and social influence," in *Psychology of group influence*, P. B. Paulus (ed.), Hillsdale, NJ: Lawrence Erlbaum. pp. 233-275

Xie, W., Li, C., Zhu, F., Lim, E. P., and Gong, X. 2012. "When a friend in twitter is a friend in life," In *Proceedings of the 4th Annual ACM Web Science Conference,* Evanston, IL, pp. 344-347.