Fragile Online Relationship:
A First Look at Unfollow Dynamics in Twitter

Workshop on Privacy and Security on Social Media
2011.1.18.
UFMG

Haewoon Kwak, Hyunwoo Chun, Changhyun Lee
Hosung Park, Sue Moon

Department of Computer Science
KAIST
Overview

• Where am I from?
• Introduction on Twitter
• Analysis of “unfollow” dynamics
• Future directions
Korea?
KAIST in Seoul? No.

INCHENON

SEUL

2hr by car, 1hr by KTX

DAEJEON

from maps.naver.com
KAIST across Gap-chon

My Office

by Yoo Chung
KAIST in snow
Frozen Korea This Week
Brief Overview of KAIST

• Founded in 1971 as the only school under the Ministry of Science and Technology (MoST)
• Has 400+ faculty, 7000+ students
  – Colleges
    • Natural Science
    • Life Science and Bioengineering
    • Engineering
    • Information Science and Technology
    • Business
  – Graduate Schools in
    • Culture Technology, Web Science and Technology, Science and Technology Policy
• Has all students waived tuition/fee
Contributions to Korean Economy

• 20% of Ph.D.s in Korean industries are KAIST graduates.
• 25% of researchers in SAMSUNG Semiconductor are KAIST graduates.
• 50% of researchers in SAMSUNG flash memory team are KAIST graduates.
• About 360 start-up companies by KAIST graduates as of 2008
  - Particularly strong in Internet services and entertainment.
Computer Science Department

- 1st graduate program in CS in Korea (1973~)
- 50 faculty, 400+ students
- The only IEEE/ACM fellow in Computer Science in Korea
- Recruiting students and faculty aggressively
Introduction on Twitter
Twitter is not a social network
Twitter is transforming the nature of news.
Twitter, a microblog service

write a short message
Twitter, a microblog service

read others' tweets
Twitter, a microblog service

check out trending topics
In most OSN

“We are friends.”
In Twitter

“"I follow you.""
"Unlike most social networks, following on Twitter is not mutual. Someone who thinks you're interesting can follow you, and you don't have to approve, or follow back."

http://help.twitter.com/entries/14019-what-is-following
http://twitpic.com/135xa – There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.
Typhoon Gonpas

http://twitpic.com/2kcw4s
8:23 AM Sep 2nd via web

http://twitpic.com/2kcnxw
9:08 AM Sep 2nd via Chromed Bird
Twitter as a News Medium

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
3. A few users reach large audience directly
4. Most users can reach large audience by word-of-mouth (WOM) quickly
media [mee-dee-uuh]

1. a pl. of medium
2. the means of communication, as radio and television, newspapers, and magazines, that reach or influence people widely

http://dictionary.reference.com/
Data collection (09/6/1~9/24)

• 41.7M user profiles (near-complete at that time)
• 1.47B following relations *publicly available
• 4262 trending topics
• 106M tweets mentioning trending topics
  – Spam tweets removed by CleanTweets
Recent studies

• Ranking methodologies [WSDM’10]
• Predicting movie profits [HYPERTEXT’10]
• Recommending users [CHI’10 microblogging]
• Detecting real time events [WWW’10]
• 4 papers with Twitter in title at WWW ’10
• The ‘entire’ Twittersphere unexplored
Finding I

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
3. A few users reach large audience directly
4. Most users can reach large audience by WOM* quickly
Why do people follow others?

• Reflection of offline social relationships

otherwise,

• Subscription to others’ messages
Sociologists’ answer

“Reciprocal interactions pervade every relation of primitive life and in all social systems”
Is following reciprocal?

- Only **22.1%** of user pairs follow each other
- Much lower than
  - 68% on Flickr
  - 84% on Yahoo! 360
  - 77% on Cyworld guestbook messages
Low reciprocity of following

- Following is not used similarly as friend in OSNs
  - Not reflection of offline social relationships

- Active subscription of tweets!
Finding II

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
3. A few users reach large audience directly
4. Most users can reach large audience by WOM* quickly
Dynamically changing trends
Majority of topics are headline news

31.5% “ephemeral”
6.9% “backintheday”
54.3% “headline news”
7.3% “persistent news”
31.5% “lynn harris”
6.9% “#redsox”

Finding III.

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
3. A few users reach large audience directly
4. Most users can reach large audience by WOM* quickly
5. Rank users based on "effective readership"
How many followers does a user have?
Are those w/ many followers active?
Finding IV.

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
3. A few users reach large audience directly
4. Most users can reach large audience by word-of-mouth quickly
In Twitter

Information

Following
Average path length: 4.1
Retweet (RT)

- Relay tweets from a following to followers
Retweet (RT)

- Relay tweets from a following to followers
Retweet (RT)

- Relay tweets from a following to followers
Retweet (RT)

- Relay tweets from a following to followers

RT @node0 Free Coffee!!!
Retweet (RT)

- Relay tweets from a following to followers

RT @node0 Free Coffee!!!

RT @node0 Free Coffee!!!

node0

node1

node2

node3

node4

node5
Retweet (RT)

- Relay tweets from a following to followers
Retweet (RT)

• Not only 1 hop neighbors
Retweet (RT)

2 hop neighbors

node0
node1
node2
node3
node4
node5
node6
We construct RT tree

- A tree with writer and retweeter(s)
Height of RT trees

1

1

2
Empirical RT trees
96% of RT trees = Height 1
Additional readers

2 additional readers by retweeter

3 followers

node0
node1
node2
node3
node4
node5
node6

Advanced Networking LAB
Boosting audience by RT
Same # of Tweets, Different Patterns of Diffusion
Two Sides of Relationship

• Formation and Dissolution
  – Formation traditionally well studied
  – Dissolution hardly much
• Why?
  – Hard to obtain data
• Proxy for dissolution
  – No exchange of email [Kossinet09]
Previous Works on Dissolution

• Half of adult friendships change over 7yr [Mollenhorst09]

• Younger children have a higher retention rate [Cairns95, Hallinan87]

• Homophily adds stability in off-line relationship [Tuma79, Kossinets09]
Two Questions We Raise

• How prevalent is unfollow?

• Why do people unfollow?
Data collection (as of September 2010)

• Korean users defined as someone who
  – Wrote a tweet in Korea
  – Keeps bio info in Korean
  – Is located in Korea
  – Has a Korean screen name

• Daily snapshots of following relationships
  – G(I): June 265th to July 15th, 2010
  – G(II): August 2nd to August 31st, 2010
Four Types of Tweet

• Tweet
  
PSSM is now starting!

• Reply
  
@Virgilio Fantastic Workshop! Thanks for having me!

• Mention
  
I am attending PSSM organized by @Virgilio and @PK!

• Retweet
  
At UFMG till tomorrow! RT @Virgilio PSSM is now starting!
Proportion of Tweet Types

Users become more informational than interactive as the number of followees increases.
How Prevalent Is Unfollow?
Follows and Unfollows

Unfollow is prevalent!
Unfollow frequent

• Mostly singular
  – 66% of unfollows are the only unfollow of the day
• But often clustered
  – 10% with 5 or more other unfollows
• On average
  – 90% of time intervals between days of unfollow is less than 9 days
Communication partner

- Reciprocal and interactive users
  - Exchange of a mention, a reply, or a retweet and vice versa
#Comm Partners vs. #Followees
Passive Nature of *Follow*

- 85.6% relationships involve no activity
- 96.3% involve 3 or fewer
- Who unfollows?
  - Remove 85.6% of no activity and among those with any activity unfollowed relationships involves less activity than unbroken relationships
Unfollow ratio vs. ego-centric ordering of relationship establishments
# Followees vs. # Unfollowees
More Retweets/Favorites
Less Likely to Be Unfollowed
The overlap of relationships vs. unfollow ratio
Why Do People Unfollow?
Interviews

Q1: Why a participant decided to unfollow.
Q2: Whether s/he thought the unfollowee was aware of being unfollowed.
Q3: If s/he broke off on other OSNs. Difference?
Q4: If s/he followed corporate accounts.
Q5: Choose 10 users s/he would never unfollow
Demographics of 22 interviewees

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>27.3</td>
<td>27</td>
<td>22</td>
<td>36</td>
<td>3.7</td>
</tr>
<tr>
<td>Favorites</td>
<td>80.7</td>
<td>1</td>
<td>0</td>
<td>851</td>
<td>199.0</td>
</tr>
<tr>
<td>Followers</td>
<td>846.7</td>
<td>164.5</td>
<td>5</td>
<td>8,772</td>
<td>2,053.9</td>
</tr>
<tr>
<td>Followings</td>
<td>600.4</td>
<td>144.5</td>
<td>5</td>
<td>7,103</td>
<td>1,562.7</td>
</tr>
<tr>
<td>Tweets</td>
<td>3,325.8</td>
<td>583.5</td>
<td>5</td>
<td>30,639</td>
<td>7,220.5</td>
</tr>
<tr>
<td>Registered days</td>
<td>449.2</td>
<td>471</td>
<td>14</td>
<td>766</td>
<td>179.1</td>
</tr>
</tbody>
</table>
Q1: Motivations behind Unfollow

• Burst (39)
  – Burst-only (13), Uninteresting topic (10), Mundane details (6), Automatically generated (4), Conversation (2), Politics (2), Different Views (1), Complains (1)
Q2: Awareness of Being Unfollowed

- A half of respondents stated that they thought unfollowees were aware of being unfollowed.
  - They did not know unfollowees in person
  - They got used to unfollowing
  - Unfollow was easy
- The other half
  - Unfollowees had too many followers to notice
  - No convenient interface to track it
  - They did not track themselves
Q3: Break-up on other OSNs?

• Not common
Q4: Corporate Accounts

• 8 out of 22 follow corporate accounts
  – 5 kept following
  – Motivation = expectation of prize winning
  – They didn’t mind occasional ad tweets, but unfollow if ads come in bursts
  – Some only participate if all participants received a gift
Q5: Whom *Not* to Unfollow?

- Most respondents chose intimate friends
- Some chose their role models
Future Work

• Break-ups in OSNs
  – Implications on cognitive and social capacity of human beings

• Security on OSNs
  – Exploit community structures for security
Q&A
Research Topics at AN Lab

• Online Social Networks

• Network Performance Measurement and Analysis

• PC-based Platforms for Future Internet
Online Social Networks

• Characteristics of OSNs

• Communities on OSN
  – Community mining (IMC ‘09, COMSNET ‘10)
  – Community evolution (in prep)

• User Interaction on OSNs
  – Declared friendship vs interaction (IMC ‘08)
  – Twitter as a news media (WWW ‘10)
  – Unfollow behavior (CHI ‘11)
  – User ranking based on retweets and human memory decay (in submission)
  – Information spreading on Twitter (in prep)
Network Performance Measurement and Analysis

- Annual Joint Workshop
- Internet-wide Path and Delay Estimation
  We can show past history of e2e delay evolution and answer what-if scenarios of AS path change
- Persistent congestion:

Intrigued?
S/W Platforms for Future Internet

• Mobile PlanetLab
  – Control channel over 3G/3.5G cellular link

• PacketShader (SIGCOMM ‘10)
  – 40Gbps PC-based S/W Router

• SSLShader (NSDI ‘11)
  – SSL Accelerator w/ GPU and AES-NI
Collaborators

• Past
  – AT&T, Telefonica I+D, U. of Cambridge
• Online Social Networks
  – Hawoong Jeong (Dept of Physics, KAIST)
  – Wonjae Lee (Institute for Social Development and Policy Research, Seoul National University)
• Network Performance Measurement and Analysis
  – Kave Salamatian (U of Savoie, France)
  – Timur Friedman, Serge Fdida (UPMC/LIP6)
  – Kenjiro Cho (WIDE, Japan)
• S/W Platforms for Future Internet
  – KyoungSoo Park (Dept of EE, KAIST)
  – Seungyeop Han (grad student, UW)