Analysis of Twitter unfollow: How often and why

SocInfo 2011
Singapore Management University

Haewoon Kwak
Hyunwoo Chun
Wonjae Lee
Sue Moon

KAIST Computer Science
KAIST

Two basic processes in network evolution

Building a relationship

Breaking a relationship

Building a relationship

Breaking a relationship

Lack of relationship breakup data

- Can you capture the breakup from
  - A paper coauthorship network?
  - A mobile phone call network?
  - An e-mail network?
  - A wall message network?

Very hard to define and capture breakup

Canceling the friendship online?

Unfortunately (?),
people do take online etiquette
seriously.

Click on a friend to begin the sacrifice.

Adam Gilad

Adam Sachs

Adam Zand

Adriana C.

Alex Argrow

Alex Mehr

Alex Panelli

Amanda Tsinonis-

Amanda Wheeler

Amiee Parco

Amy Alex Andersen

Amy Canaday

▶ Share The Sacrifice

FRIENDSHIP IS STRONG, BUT THE **WHOPPER®** IS STRONGER.

Use this application and be rewarded with a free flame-broiled WHOPPER® when you sacrifice 10 of your Facebook friends. Each friend will be notified so choose wisely.

BURGER KING

A proxy for relationship breakup

- Disappearance of e-mail exchange

"Alice & Bob had exchanged e-mails frequently.

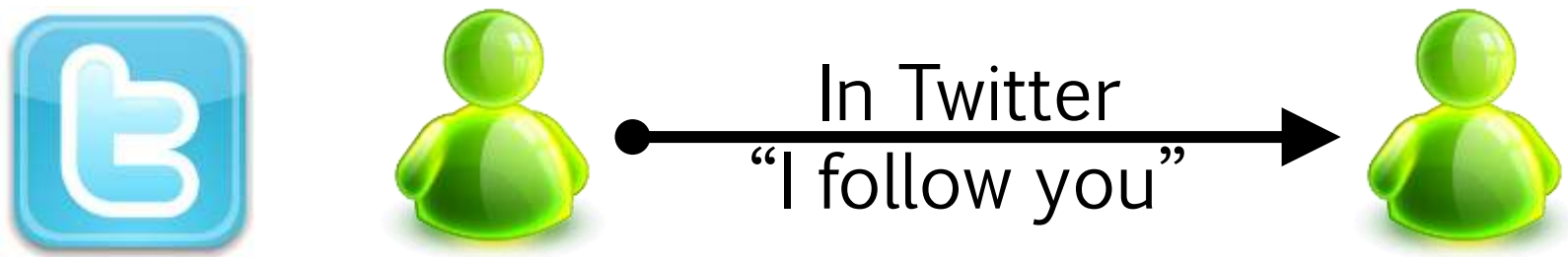
At some point they didn't do any more."

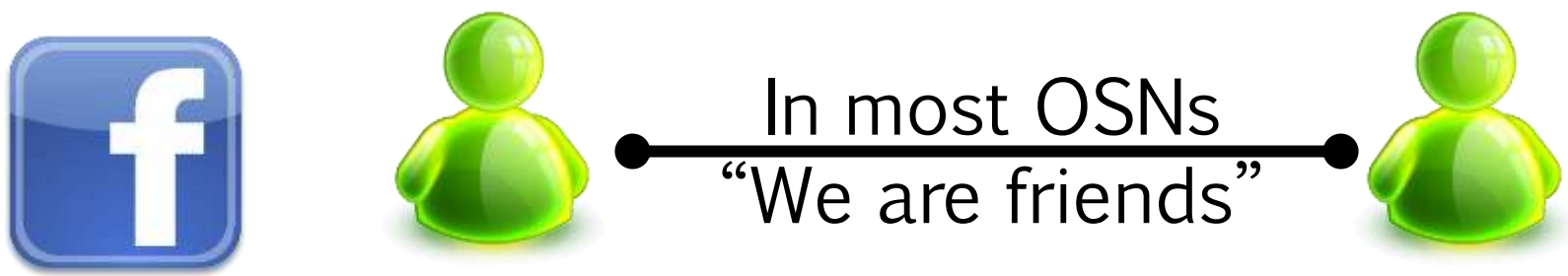
But they were doing instant messaging.

A proxy is not always accurate!

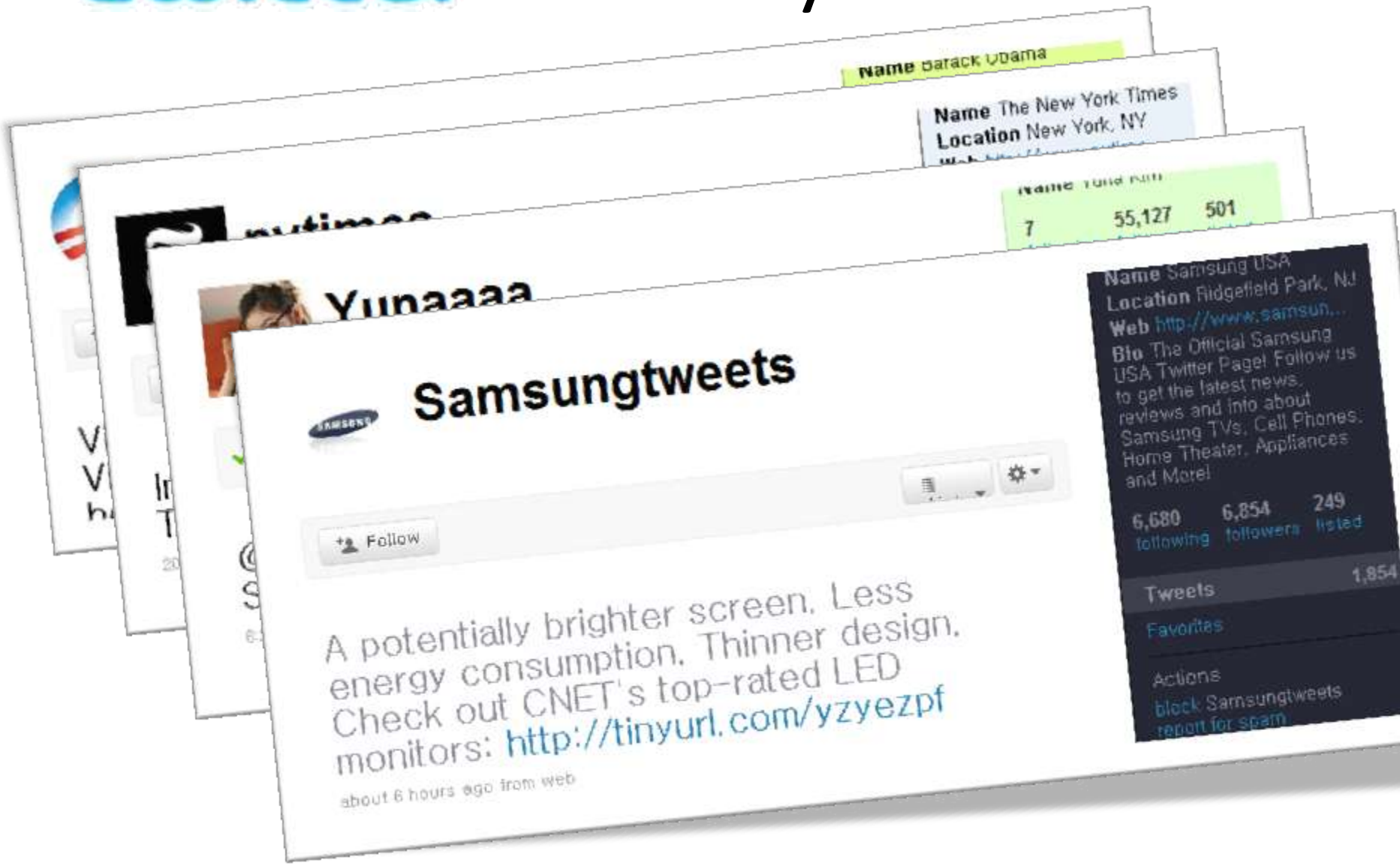
How about Twitter?

Different from other OSNs





twitter for everyone



jkrums
Follow
http://twitpic.com/135xa – There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.
US AIRWAYS

www.slideshare.net/haewoon/what-is-twitter-a-social-network-or-a-news-media-3922095

t is Twitter, a Social Network or a News Media? 1 year ago

Email Favorite Download Embed

What is Twitter,
a Social Network or a News Media?

Haewoon Kwak Changhyun Lee Hosung Park Sue Moon
Department of Computer Science, KAIST, Korea

19th International World Wide Web Conference (WWW2010)

Haewoon Kw
67523 views, 2

Related More

Unfollow in Twitter

- Intentional action to break a relationship
- No need for an approval
- No notification

  to the unfollowed

Google Calendar | Twitter / Home

twitter.com

twitter | Search | Home | Profile | Messages | Who To Follow

Read others' tweets in my timeline

Write a short message

Check out trending topics

Timeline | @sbmoon

gihong Gihong
페친께서 진행
http://j.mp/p3CF5k
6 minutes ago

TechCrunch TechCrunch
Travel Much? The Unlocked iPhone 4S Will
November tcrn.ch/qp0VyP
10 minutes ago

mooseang micky kim (a.k.a ㅎㅅ)
슬레이어 레인인블러드 앨범 발매 25주년 되는 날이라는군요. 기념으로 한 번 들어줘야겠네요. ㅋㅋㅋ
37 minutes ago
Retweeted by ntrolls

TechCrunch TechCrunch
ZipPay To Launch New Mobile Payments Service Cheaper Than Square tcrn.ch/rirKky by @sarahintampa
35 minutes ago

wwoo_ct Woontack Woo (우운택)
#AR #CAMAR for Ubiquitous VR is out! bit.ly/lRZghL ▸ Top stories today via @outdoor_ads @bdwassom @uvisor_jobs_ar @carstufffeed

Your Tweets 2,027
1 hour ago : SocInfo chairs shared Pol

Following 170 | Follower

Who to follow · refresh · view all
SFGiants San Francisco Giants
Promoted
Follow

Trends: Worldwide · change
Paranormal Activity 3 Promoted
#Top10Lies
#3Palabras
#СПАСИБОПУТИНУЗАЭТО
SELENA GOMEZ IS GORGEOUS
Y U SO FIT
MARVIN IS FLAWLESS
ASTON IS SEXY
HURRICANE THE BOOK
ZAYN ROGAN JOSH MALIK

Four Types of Tweets

Tweet

> Last Day of SocInfo!

Reply

> @EePeng Thanks for having me!

Mention

> I am attending SocInfo 2011 organized by @EePeng

Retweet

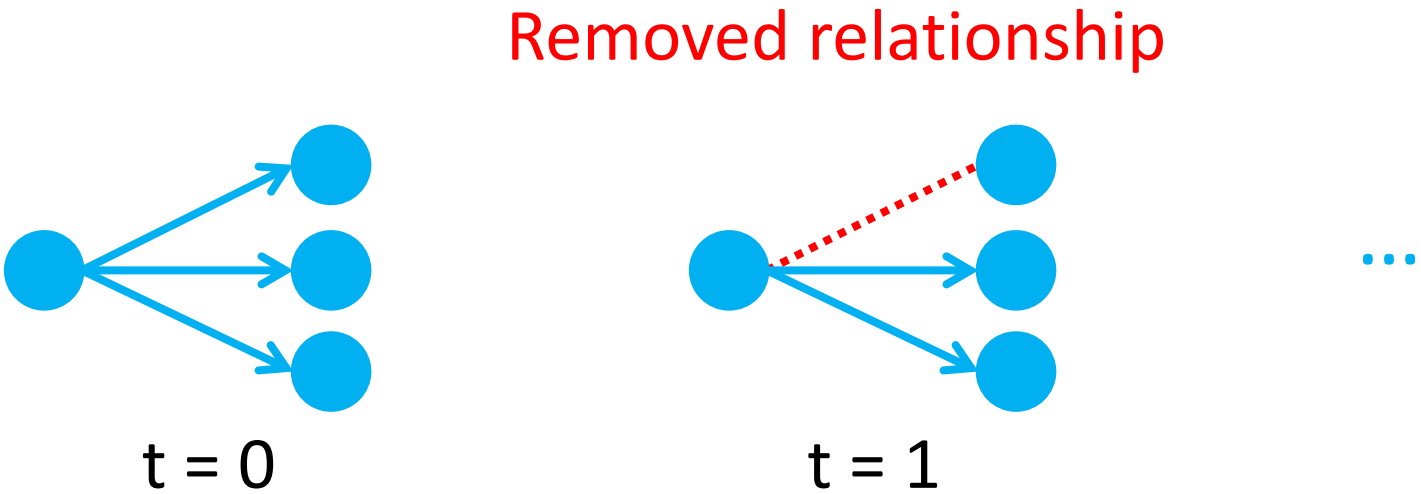
> Wow lots of fun talks and great people! RT @sbmoon Last Day of SocInfo!

Recent changes in Twitter API

Our Unfollow Study

- Macroscopic statistics
- User interviews
- Explanatory model

No official records for unfollow



- Compare two follow networks and detect removed relationships

Twitter now too big to crawl

- In 2009 it took 3 months with 20+ hosts to crawl and get the entire 40 million user profiles
- Now in 2011 it has more than 400 million users
- How to sample?

Identify a group with common culture

- Collect 1.2M Korean-speaking users identified by Korean in tweets, bio, location, or name

Data collection

- Collect daily snapshots of follow networks
  - G(I): June 25th to July 15th, 2010
  - G(II): August 2nd to August 31st, 2010

- Time resolution = a single day

Korean follow network grew fast

- Increasing # of users
  - G(I): 718,077 → 870,057 +7,599/day
  - G(II): 956,261 → 1,203,196 +8,515/day
- Increasing (high) reciprocity
  - G(I): 56 → 58%
  - G(II): 61 → 62%
- Increasing avg. # of followees
  - 59.7 → 75.7

Macroscopic Statistics

People unfollow frequently

- 43% of active users unfollow at least once during 51 days

- Average number of unfollows per person
  - 15.4 in G(I)
  - 16.1 in G(II)

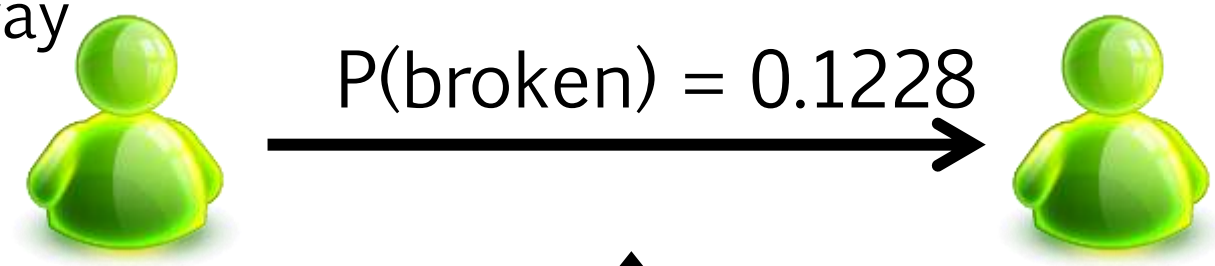
Link removal frequently occurs in a 'growing' network

Factors that correlate with unfollow

- Reciprocity of relationships
- Duration of a relationship
- Followee's informativeness
- Overlap of relationships

One-way relationships are fragile

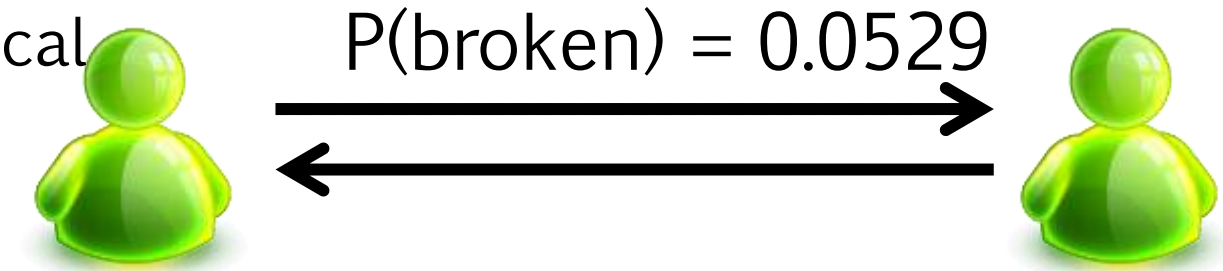
One-way



≠

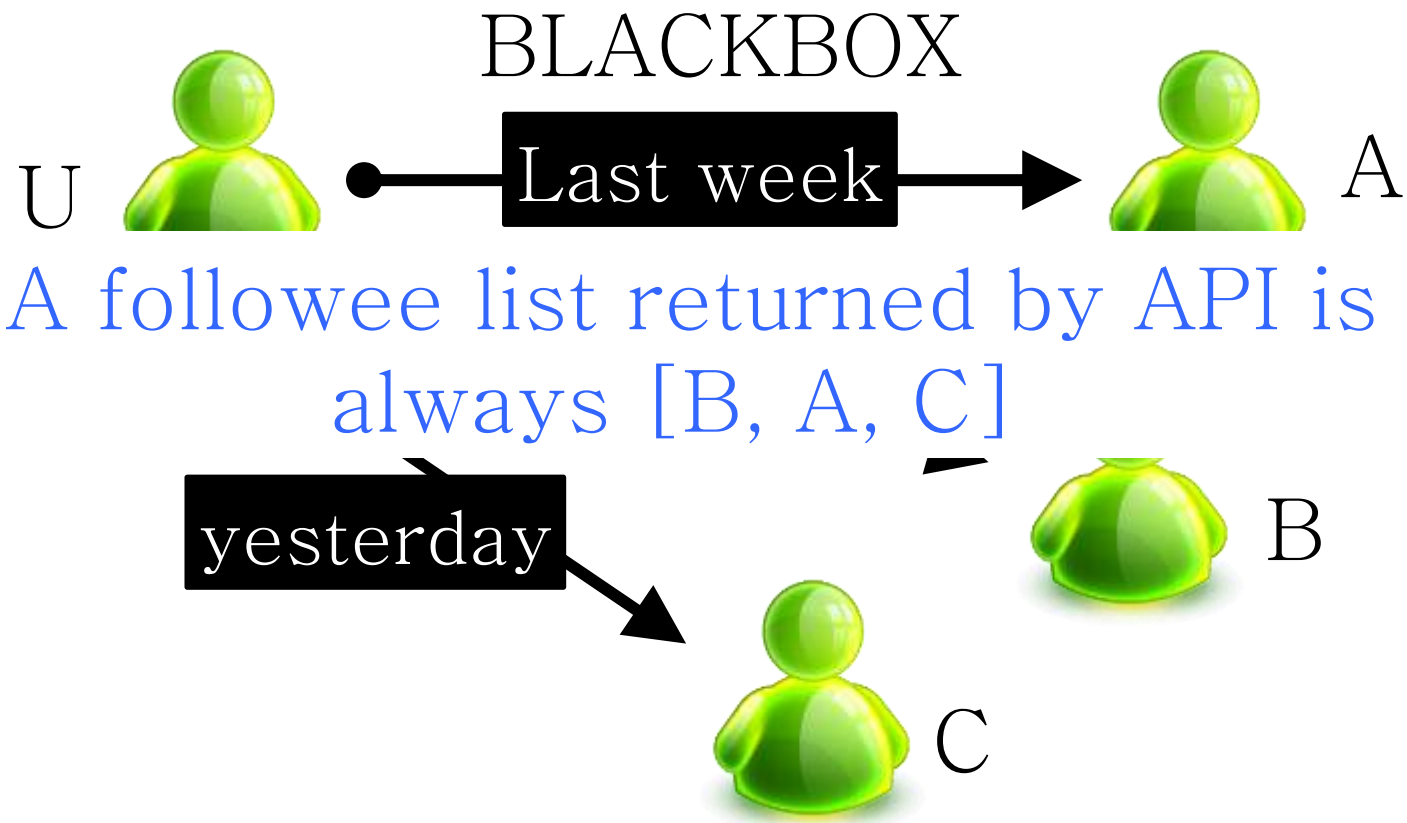
"Emotional closeness"

Reciprocal



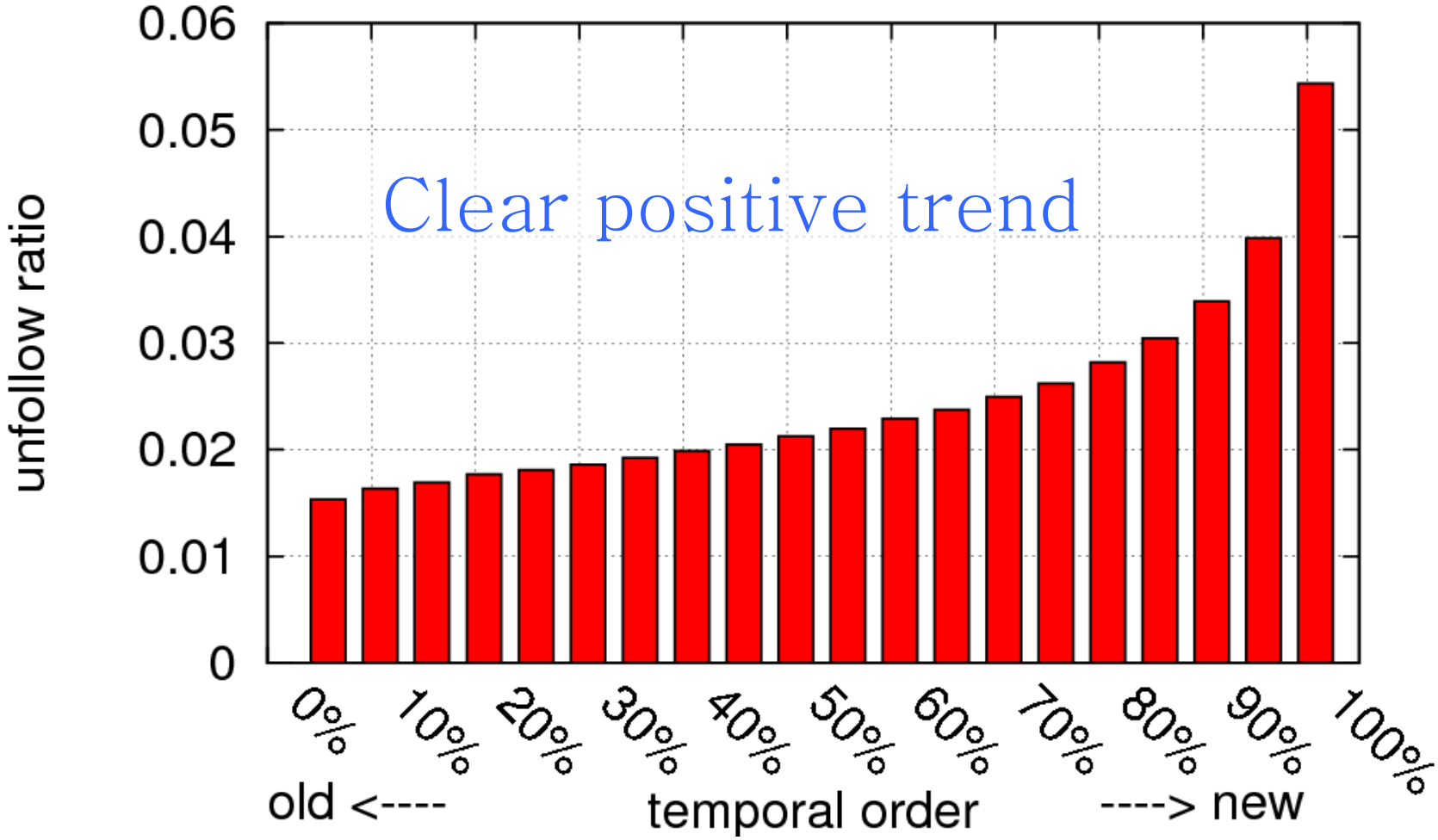
Conditional P(remaining will be broken|one is broken) = 0.2345

No knowledge when following started

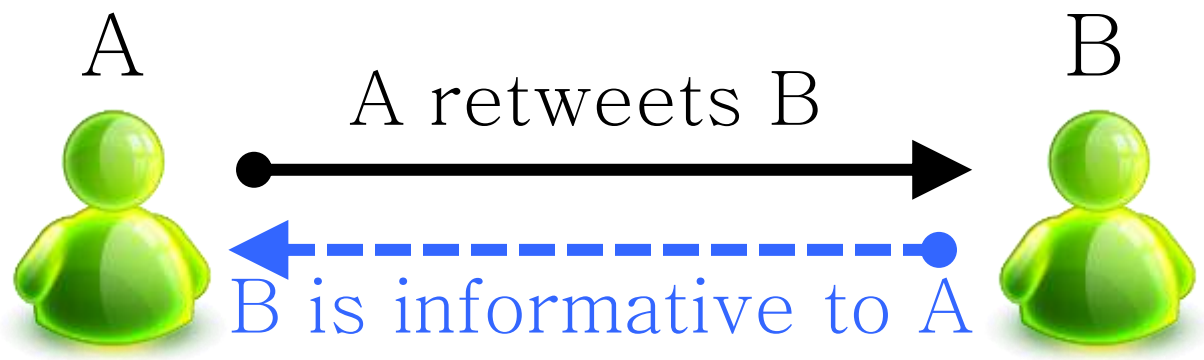


Instead, we have a temporal 'order' of relationships in a personal network

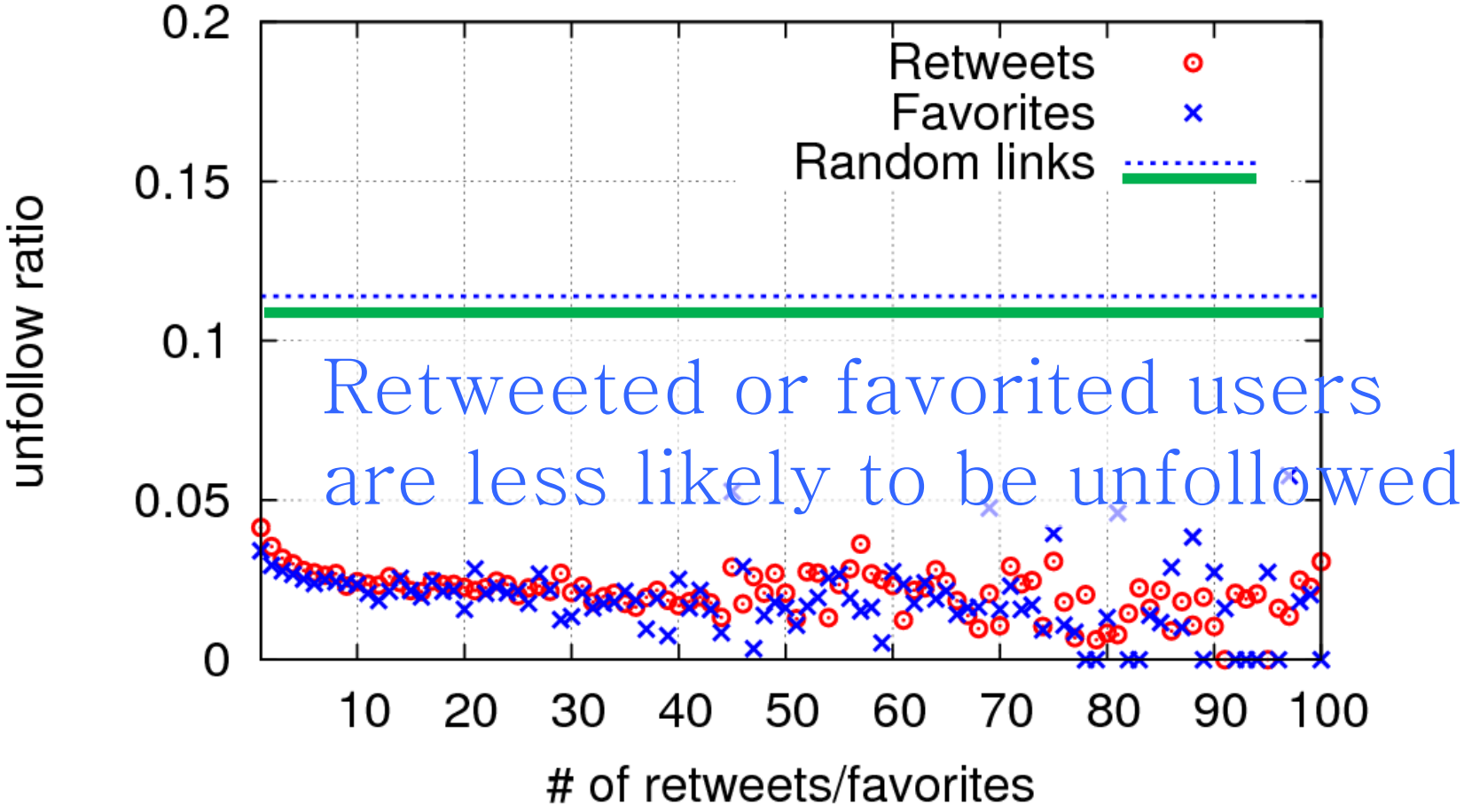
Newer relationships are more fragile



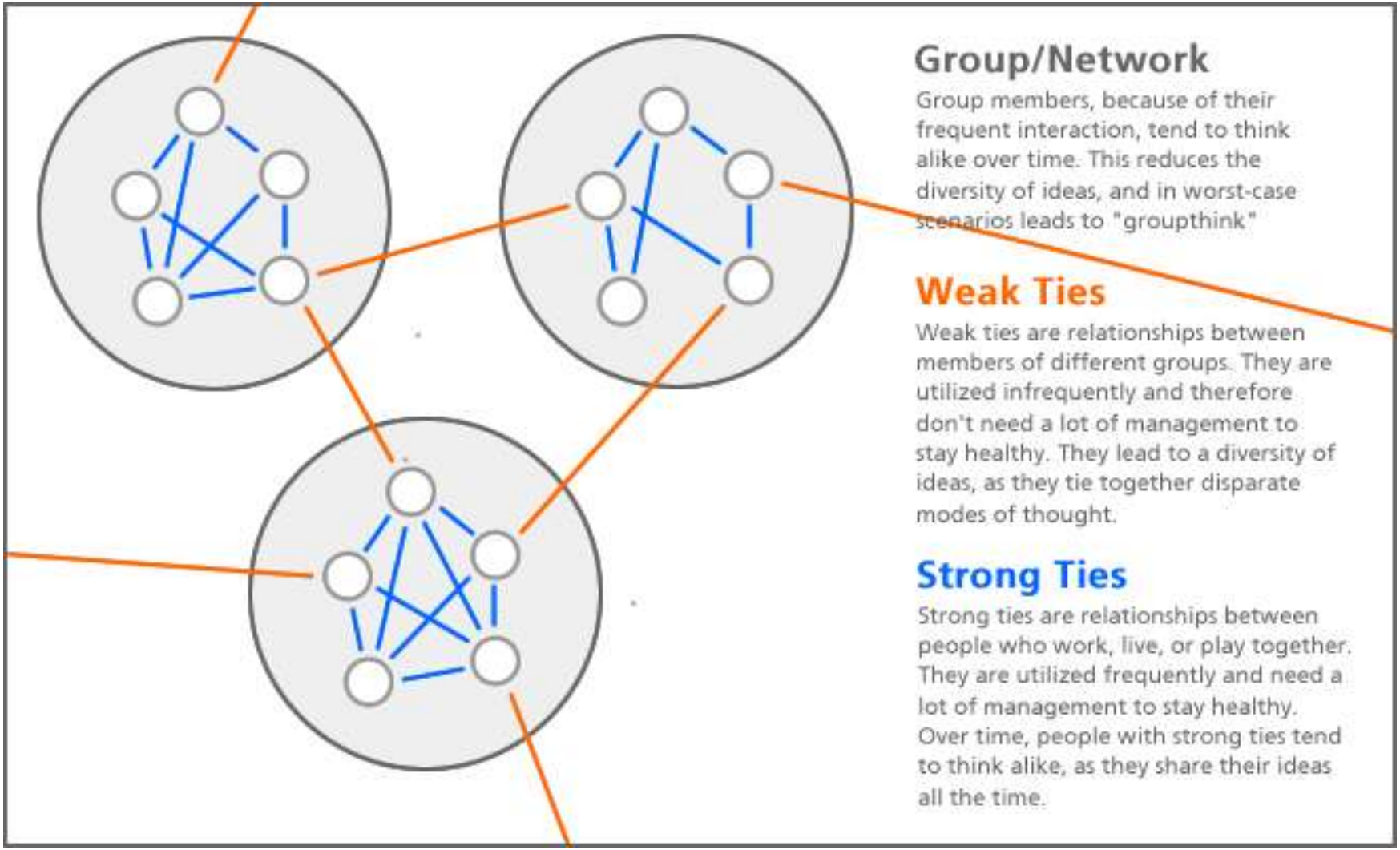
Informativeness of users


A
B
A retweets B
B is informative to A

Non-informative relationships are fragile

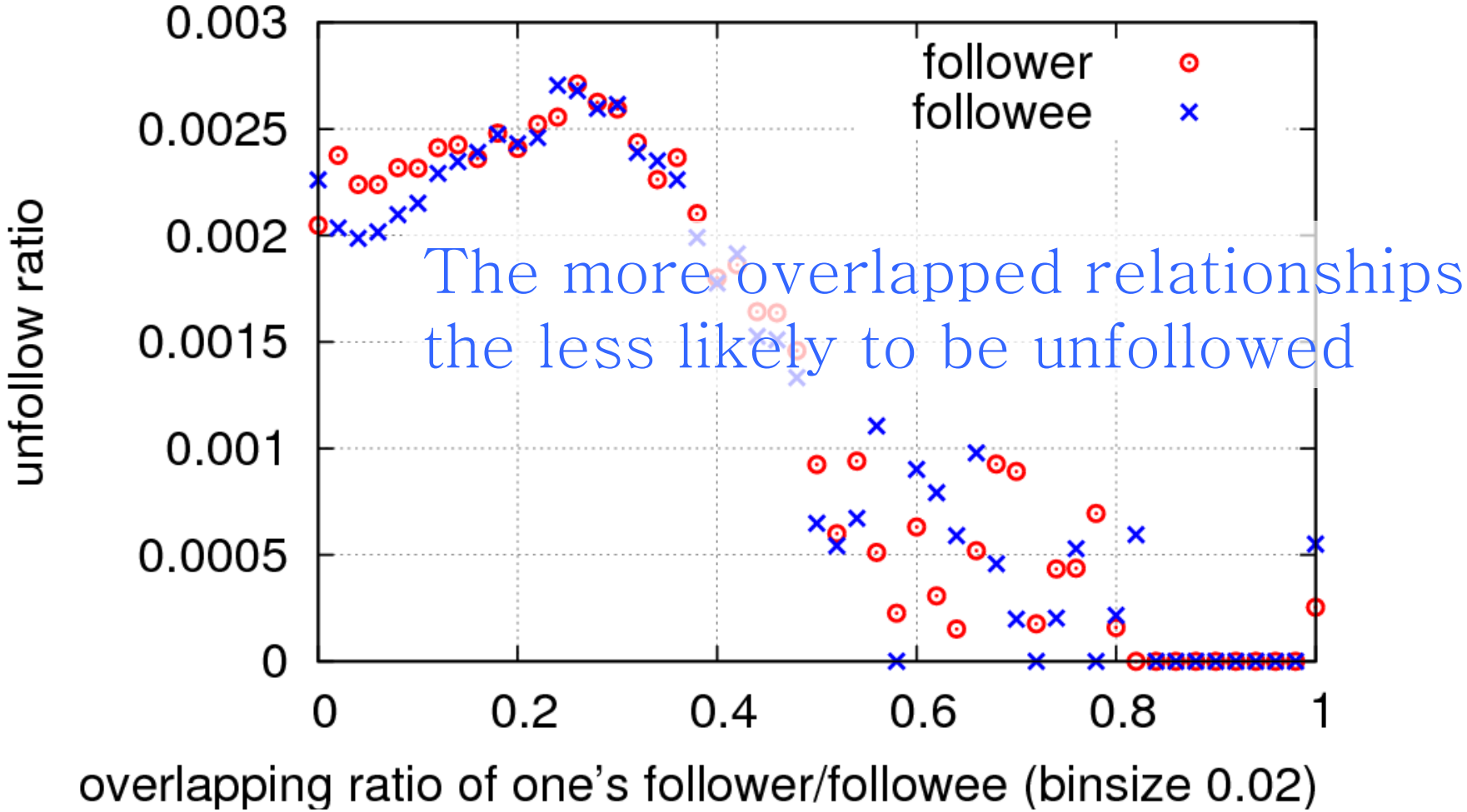


Strong ties & weak ties



http://bokardo.com/archives/weak-ties-and-diversity-in-social-networks/

Weak ties are fragile



No interaction ≠ breakup

- 85.6% of relationships do not involve any single reply, mention, or retweet
  - 96.3% involve 3 or fewer

- People just subscribe to others' tweets passively

Why our study of `unfollow' is important for the study of breakup

User Interviews

22 online & face-to-face interviews

- Recruited by word-of-mouth
- Semi-structured
- Logging & camera recording

11 Male, 11 Female

	Mean	Median	Min	Max	Std. dev.	Distribution
Age	27.3	27	22	36	3.7	
Favorites	80.7	1	0	851	199.0	
Followers	846.7	164.5	5	8,772	2,053.9	
Followings	600.4	144.5	5	7,103	1,562.7	
Tweets	3,325.8	583.5	5	30,639	7,220.5	
Registered days	449.2	471	14	766	179.1	

Top reasons in unfollow

1. Burst tweets
2. Tweets about uninteresting topics
3. Tweets about mundane details of daily life
   - Automatically generated tweets (e.g., 4sq)
4. Tweets about political issues

Explanatory Model

No explanatory model taking structural and interactional properties into consideration

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

Haewoon Kwak, Hyunwoo Chun, and Sue Moon
Department of Computer Science, KAIST
291 Daehak-ro, Yuseong-gu, Daejeon, Korea
{haewoon, hyunwoo}@an.kaist.ac.kr, shmoon@kaist.edu

ABSTRACT
We analyze the dynamics of the behavior known as 'unfollow' in Twitter. We collected daily snapshots of the online relationships of 1.2 million Korean-speaking users for 51 days as well as all of their tweets. We found that Twitter users frequently unfollow. We then discover the major factors, including the reciprocity of the relationships, the duration of a relationship, the followees' informativeness, and the overlap of the relationships, which affect the decision to unfollow. We conduct interview with 22 Korean respondents to supplement the quantitative results. They unfollowed those who left many tweets within a short time, created tweets about uninteresting topics, or talked about the mundane details of their lives. To the best of our knowledge, this work is the first systematic study of the behavior in Twitter.

Author Keywords
Unfollow, computer-mediated communication, Twitter, online relationship

ACM Classification Keywords
H.5.0 Information Systems: Information Interfaces and Presentation - General

General Terms
Human Factors

INTRODUCTION
Relationship formation and dissolution are two basic processes of relationship change and evolution in personal networks. Studies of relationship formation and dissolution mostly rely on surveys and interviews, both of which require considerable effort in terms of time and labour. Online social networks (OSN) aid researchers in at least two ways, such as (i) they contain a huge archive of human behavior related to online relationships, and (ii) they allow easy access. Studies of online relationship formation are straightforward, as most OSNs offer simple means of establishing online relationships, often referred to as a 'friend'. By contrast, research on the topic of online relationship dissolution has not been extensively conducted due to the lack of data; an online friend relationship remains rigid regardless of the actual relationship [28]. Researchers thus use proxies to represent the state of relationship dissolution. For example, a study of relationship dissolution in email networks assumes that the disappearance of online activities (the exchange of emails) reflects this type of dissolution [16]. However, a disappearance of communication cannot be directly translated to the dissolution of a relationship in most cases, because it is difficult to capture all communication means [23] and to regard the absence of an event as strictly intentional. The key is [illegible] relationship.

Twitter has been redefining human behavior [illegible] features of Twitter, in contrast to the reciprocal friendship in most other OSNs, e.g., friends in Facebook. Literally, neither an invitation [illegible] following; people can follow [illegible] user logs into Twitter [illegible] follows (followees). The stream of followees' tweets is called a *timeline*. A user can easily stop following (*unfollow*) and needs no confirmation from the followee to do so. Unfollow, thus, is not a proxy but a verifiable action of breaking an online relationship. In the rest of the paper, we use unfollow both as a noun and a verb.

In this work, we analyze the dynamics of the unfollow behavior to understand online relationship dissolution. The two research questions explored here are: (i) what are the characteristics of the unfollow behavior? and (ii) why do people unfollow others? To address the first research question, we collected daily snapshots of the follow relationships of 1.2 million Korean-speaking users over the course of 51 days as well as their tweets. By comparing the daily snapshots, we confirm that unfollow is prevalent in Twitter. We have found that the reciprocity of the relationship, the duration of the relationships, the followees' informativeness, and the overlap of relationships are critical in the decision to un-

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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
Copyright 2011 ACM 978-1-4503-0267-8/11/05...$10.00.

The Impact of Network Structure on Breaking Ties in Online Social Networks: Unfollowing on Twitter

Funda Kivran-Swaine, Priya Govindan and Mor Naaman
Rutgers University
funda@rutgers.edu, priyagn@cs.rutgers.edu, mor@rutgers.edu

ABSTRACT
We investigate the breaking of ties between individuals in the online social network of Twitter, a hugely popular social media service. Building on sociology concepts such as strength of ties, embeddedness, and status, we explore how network structure alone influences *tie breaks* – the common phenomena of an individual ceasing to "follow" another in Twitter's directed social network. We examine these relationships using a dataset of 245,586 Twitter "follow" edges, and the persistence of these edges after nine months. We show that structural properties of individuals and dyads at Time 1 have a significant effect on the existence of edges at Time 2, and connect these findings to the social theories that motivated the study.

ACM Classification Keywords
H.5.0 [Information interfaces and presentation]: General

INTRODUCTION
Many of today's popular Web applications and social media services, like Facebook, Twitter, or LinkedIn, are built on [illegible] application. [illegible] user's activity [illegible] streams [12] like Twitter and Facebook see the stream of updates from their contacts when they log in). These contacts reflect various types of relationships, including friendship, kinship, common interests, attention, or information exchange. The aggregates of contact lists in each of these services result in immense, articulated *online social networks*. As these services shift the communication and information fabric of our society, the dynamics of the networks they support are important to understand and reason about in depth.

In particular, the articulated online social networks in these services change and evolve as individuals form new ties, or break existing ties to others. These structural changes, observed over time, are far from random, and depend on various factors that affect the relationship between the users. Most work in the Computer Science and HCI fields has focused on the dynamics and models of tie *creation* [5,9,13]. In particular, researchers considered the *structural* aspects of the social network that predict formation of ties [5,10,15]. Here, we also focus on the social network structure, but examine breaking and persistence of existing ties, rather than tie creation.

The topic of breaking and persistence of ties is exceedingly important. Tie breaks impact the dynamics and activity in online services over time, and, as we show below, are common on services like Twitter. Moreover, the act of "de-[illegible] in online social systems might impact social [illegible] within and beyond the online world, in part due [illegible] articulated nature of the network, where the ties can be observed, and their nonexistence could be detected.

While a number of sociology studies have examined tie [illegible] the topic has not received [illegible] field [1]. This gap is [illegible] data, or the fact that unlike online ties, social ties often decay rather than "formally" break, making measurement of the phenomena harder. The initial findings from sociology research, however, suggest that the structural aspects of individuals' social networks are strong determinants for persistence and decay of ties [1,11]. We now have an opportunity to examine these theories in online social networks, and at a scale that was not available before. The dynamics of these online networks may help us, then, to reason about other types of networks and social ties, and discover patterns that may shed light on social phenomena in different contexts.

We use two snapshots of the Twitter social network to study the breaking of ties. Twitter is founded on an articulated online network, and allows users to read updates from others that they "follow". Thus, users create a *directed* social network that reflects attention and transfer of information [5,7,12,13]. We borrow from sociology theory to frame our investigation of tie breaks and network structures using the concepts of tie strength, embeddedness as well as power and status. Our largescale analysis uses information from 245,586 dyadic relationships on Twitter at Time 1, and the persistence or break of edges in these dyads by Time 2. The analysis aims to answer the following research question: *What structural social network properties of nodes and dyads predict the breaking of ties?*

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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
Copyright 2011 ACM 978-1-4503-0267-8/11/05...$10.00.

Interactional properties describe dynamics

- Has a followee sent a reply to a follower?
- Has a user mentioned a follower in any of one's tweet?
- Do a user and a follower share common topics of interest?

Sociological concepts in persisting ties

Homophily
Reciprocal services
Tie strength
Power and prestige

+ Twitter-specific feature: Informativeness

Multiple logistic regression model

- Binary dependent variable
  - Whether a relationship at t_0 will be broken (1) or persisted (0) at t_1

- 78 Independent variables
  - From structural properties
    - # of followers, # of followees, # of common followers, …
  - From interactional properties
    - # of replies, # of retweets, # of communication partners, …

Filtered variables

- Removing multicollinearity
  - # of common followers & followees & neighbors
  - # of follower & those who reply to ego
  - # of those who exchange replies & all replies
  - # of received replies & mentions
  - …

- We filter out 36 variables and 42 remained
- With stepwise regression, we further winnowed down to 39 variables

	Dependent variables	Coefficient	Odds ratio
	Structure properties		
u	Followees	-4.13e-04 (1.50e-06)	1.000***
u	Followers	5.59e-05 (8.81e-07)	1.000***
u	Followees/Followers	-1.71e-03 (1.98e-05)	0.998***
u	**Follow-back ratio**	-0.21 (4.02e-03)	0.810***
f	followees	2.67e-05 (2.53e-07)	1.000***
f	followers	-2.68e-06 (8.01e-08)	1.000***
f	Followees/Followers	1.68e-03 (1.28e-04)	1.002***
f	**Follow-back ratio**	0.82 (4.68e-3)	2.282***
$u \rightarrow f$	Order of follow	-1.57e-08 (8.41e-11)	1.000***
$u \rightarrow f$	^{N}Order of follow	2.58e-08 (6.71e-10)	1.000***
$u \leftrightarrow f$	Prestige	-1.22e-06 (2.37e-07)	1.000***
$u \leftrightarrow f$	**Reciprocity of follows**	-0.80 (2.50e-03)	0.451***
$u \leftrightarrow f$	Common followees	-7.40e-05 (9.12e-06)	1.000***
$u \leftrightarrow f$	^{N}**Common followees**	-2.89 (2.78e-02)	0.056***
	Activity properties		
u	Tweets	2.76e-04 (9.40e-07)	1.000***
u	URL tweets	-1.40e-04 (7.24e-06)	1.000***
u	Auto-generated tweets	5.26e-04 (7.32e-05)	1.001**
u	Popularity	-1.44e-04 (1.07e-06)	1.000***
f	Tweets	-4.27e-05 (1.11e-06)	1.000***
f	URL tweets	-4.32e-05 (8.74e-06)	1.000***
f	Auto-generated tweets	-2.86e-04 (1.14e-04)	1.000*
f	Popularity	-6.04e-06 (1.02e-07)	1.000***
$u \rightarrow f$	Replies	2.85e-03 (1.32e-04)	1.003***
$u \rightarrow f$	^{N}**Replies**	-0.67 (2.22e-02)	0.516***
$u \rightarrow f$	Mentions	-9.23e-04 (3.70e-04)	0.999*
$u \rightarrow f$	Retweets	-1.21e-02 (9.65e-04)	0.988***
$u \rightarrow f$	Favorites	-5.06e-02 (1.95e-03)	0.951***
$u \rightarrow f$	^{N}Favorites	-3.95e-04 (2.53e-02)	1.000***
$u \leftarrow f$	Replies	-1.58e-02 (3.91e-04)	0.984***
$u \leftarrow f$	^{N}**Replies**	-1.49 (3.52e-02)	0.225***
$u \leftarrow f$	^{N}**Mentions**	-0.66 (5.19e-02)	0.516***
$u \leftarrow f$	Retweets	-5.17e-02 (2.37e-03)	0.950***
$u \leftarrow f$	^{N}**Retweets**	-0.76 (4.92e-02)	0.469***
$u \leftarrow f$	Favorites	-6.11e-02 (5.01e-03)	0.941***
$u \leftarrow f$	^{N}**Favorite**	-0.35 (4.88e-02)	0.706***
$u \leftrightarrow f$	**Common hashtags**	-0.12 (1.29e-03)	0.883***
$u \leftrightarrow f$	^{N}**Common hashtags**	-2.2 (4.20e-02)	0.109***
$u \leftrightarrow f$	Days since first comm.	-1.95e-03 (4.8e-05)	0.998***
$u \leftrightarrow f$	Days since last comm.	-1.08e-04 (3.38e-06)	1.000***
	Removed variables		
$u \rightarrow f$	^{N}Mentions	3.49e-02 (3.46e-06)	1.000
$u \rightarrow f$	^{N}Retweets	6.25e-03 (2.28e-02)	1.000
$u \leftarrow f$	Mentions	1.09 (5.68e-04)	1.001

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

* Odds ratio is rounded to thousandths.

Interpretation of models

Emotional closeness

- If not reciprocal, 2.17x times likely to unfollow

Homophily

- If jaccard coeff. of hashtags decreases 0.2, 2.75x

Tie strength

- If overlap of followees decrease 0.2, 2.12x

Receiving actions rather than giving ones

- Whether feedback is receiving or giving is more important variable than the opposite
  - People appreciate receiving feedback
  - Consistent with the study of social capital at Facebook

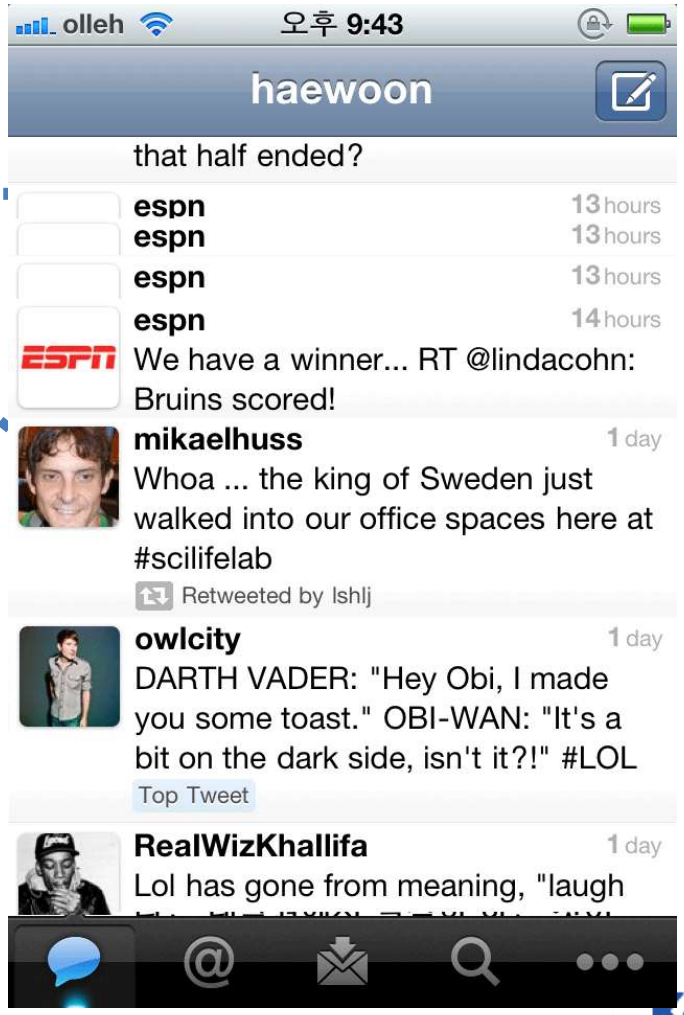
Summary

- Unfollow is a clear sign of breaking a relationship
  - No interaction is NOT a good indicator of breaking a relationship due to user's passivity in Twitter

- Quantitative and qualitative studies of unfollow
  - Reveal its characteristics and motivations
  - Discover important structural and interactional properties

Application: Practical uses

~~Burst tweets, automatically generated tweets, …~~

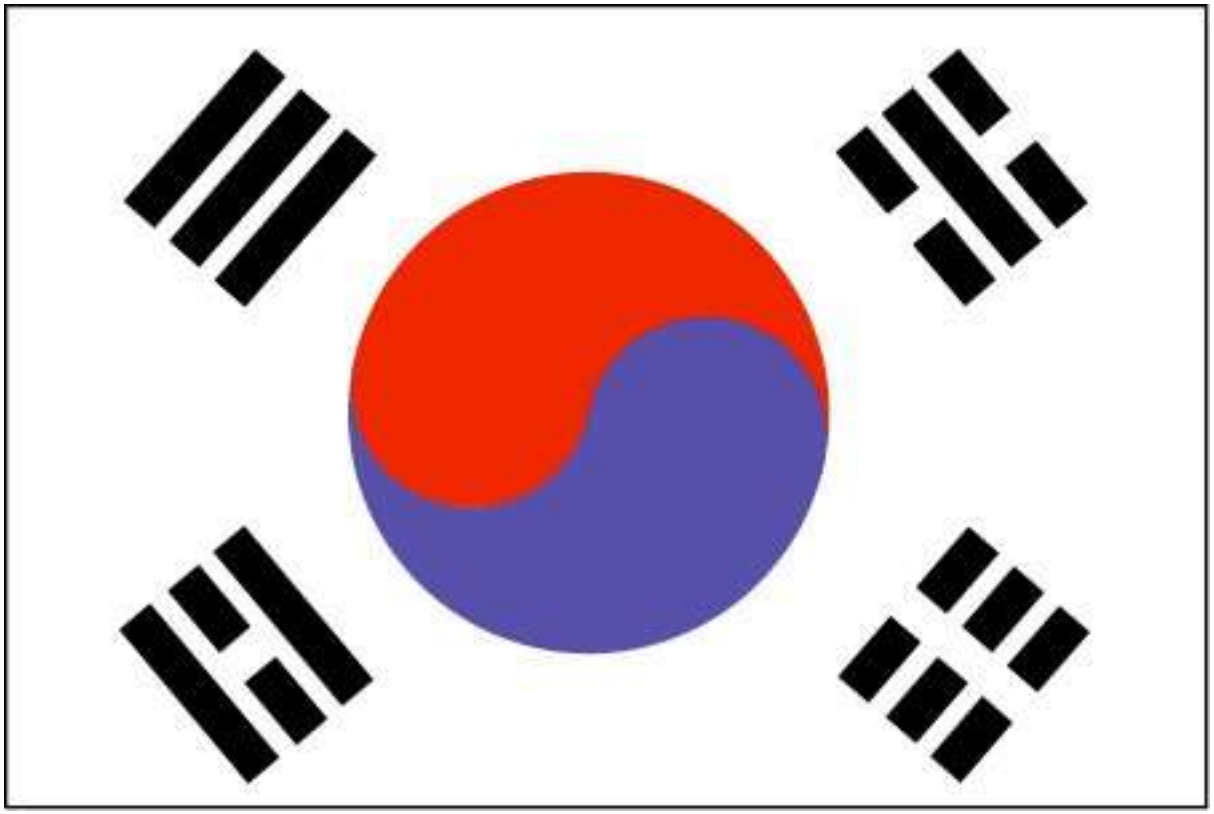


Application: Theoretical models

- Existing models for growing networks deal with link additions only

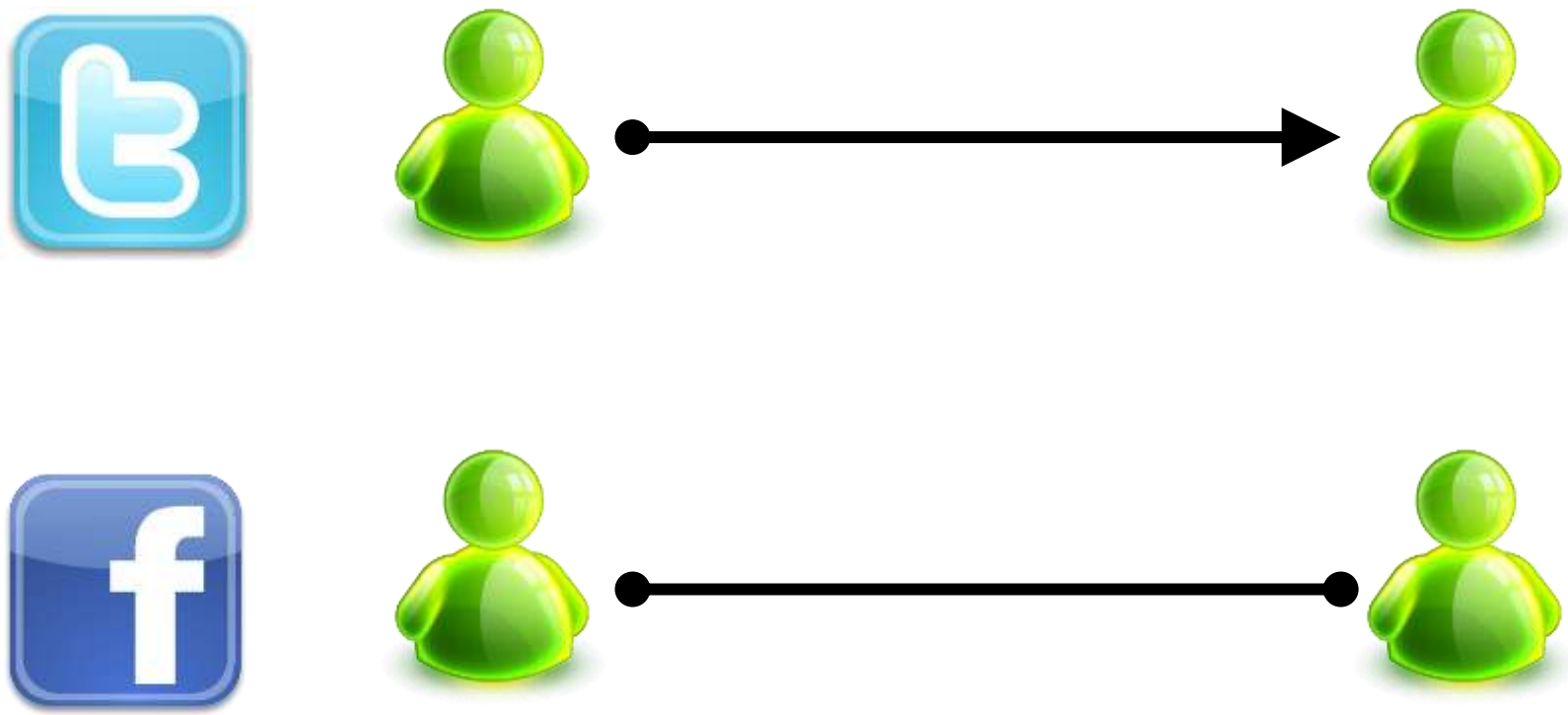
- Our studies reveal frequent link removals in a growing network

One population - Korean



How can we generalize our results?

Follow ≠ Friend



Thanks!

Questions?