Self-Disclosure and Relationship Strength in Twitter Conversations

JinYeong Bak, Suin Kim, Alice Oh
{jy.bak, suin.kim}@kaist.ac.kr, alice.oh@kaist.edu
Department of Computer Science, KAIST
In social psychology

- Degree of self-disclosure in a relationship depends on the strength of the relationship
- Strategic self-disclosure can strengthen the relationship
Introduction

In social psychology

- Degree of self-disclosure in a relationship depends on the strength of the relationship
- Strategic self-disclosure can strengthen the relationship

Hello~

Hi
In social psychology

- Degree of self-disclosure in a relationship depends on the strength of the relationship
- Strategic self-disclosure can strengthen the relationship

I'm from Korea.

I want to be your friend.

Where are you from?
Hypothesis

Twitter conversations also show a similar pattern

- Dyads with high relationship strength show more self-disclosure behavior
- Dyads with low relationship strength show less self-disclosure behavior
Methodology

- **Twitter Data**
  - 131K users
  - 2M conversations
- **Relationship Strength**
  - Chain frequency (CF)
  - Chain length (CL)
- **Self-Disclosure**
  - Personal information
  - Open communication
  - Profanity
- **Analysis with Topic Models**
  - Latent Dirichlet allocation (LDA, [Blei, JMLR 2003])
  - Aspect and sentiment unification model (ASUM, [Jo, WSDM 2011])
Twitter Conversation

- A Twitter conversation chain
  - 3 or more tweets
  - at least one reply by each user

- Our Twitter conversation data
  - Oct 2011 to Dec 2011
  - 131K users
  - 2M chains
  - 11M tweets
Relationship Strength

- Social psychology literature states relationship strength can be measured by communication frequency and length [Granovetter, 1973; Levin and Cross, 2004]

- **CF**: chain frequency
  - The number of conversational chains between the dyad averaged per month

- **CL**: chain length
  - The length of conversational chains between the dyad averaged per month

- **Relationship strength**
  - A high CF or CL for a dyad means the relationship is strong
  - A low CF or CL for a dyad means the relationship is weak
Self-Disclosure

- **Open communication - Openness**
  - Negative openness
  - Nonverbal openness
  - Emotional openness
  - Receptive openness – difficult to find in tweets
  - General-style openness – not clearly defined in the literature

- **Personal Information**
  - Personally Identifiable Information (PII)
  - Personally Embarrassing Information (PEI)

- **Profanity**
  - nigga, ass, wtf, lmao
Self-Disclosure - Openness

Negative openness

- I have definitely lost my phone...again :( :( :
- How did you lose it this time?
- I can't see it :

Method
- We use ASUM with emoticons as seed words
  [ “Aspect and sentiment unification model for online review analysis”, Jo, WSDM’11 ]
- ASUM is LDA-based joint model of topic and sentiment
- ASUM takes unannotated data and classifies each sentence (tweet) as positive/negative/neutral
Self-Disclosure - Openness

Nonverbal openness

Method

- We look for emoticons, ‘lol’, ‘xxx’
- Emoticons are like facial expressions -- :) :( :P
- ‘lol’ (laughing out loud) and ‘xxx’ (kisses) are very frequently used in a similar manner to nonverbal openness
Self-Disclosure - Openness

Emotional openness

- your having a great pregnancy that's great! How is your bf? Is he happy about the baby?
  - yeah I am. and yeah he's more happy & excited then me.

Method

- Look for tweets that contain common expressions of feeling words [We feel fine (Harris, J, 2009)]
Self-Disclosure – Personal Information

Personally Identifiable Information (PII)

- I live between Camberwell, Peckham, Brixton & Dulwich.
- I lived in Primros Hill *the Island* for fifteen years.

Ex) name, location, email address, job, social security number

Personally Embarrassing Information (PEI)

- First personal training session since hospital at 9.00 am, MRI scan at 3.30.
- Hope the scan went well.
- I’m a bit claustrophobic so I nearly pressed the alarm as soon as I went in the tube!
- I’m claustrophobic too, Good luck with the result!

Ex) clinical history, sexual life, job loss, family problem
Self-Disclosure – Personal Information

- Discover topics in each conversation
  - Use LDA [Blei, JMLR 2003] with $k = 300$
  - LDA outputs a topic proportion for each conversation
  - LDA outputs a multinomial word distribution for each topic

- Find related topics
  - Annotate conversations that best represent each topic
  - Use Amazon Mechanical Turk
  - Turkers annotated conversations for
    - existence of PII
    - existence of PEI
    - keywords
Self-Disclosure – Personal Information

Example of PII, PEI and Profanity topics

- Shown by high probability words in each topic

<table>
<thead>
<tr>
<th>PII 1</th>
<th>PII 2</th>
<th>PEI 1</th>
<th>PEI 2</th>
<th>PEI 3</th>
<th>Profanity</th>
</tr>
</thead>
<tbody>
<tr>
<td>san</td>
<td>tonight</td>
<td>pants</td>
<td>teeth</td>
<td>family</td>
<td>nigga</td>
</tr>
<tr>
<td>live</td>
<td>time</td>
<td>wear</td>
<td>doctor</td>
<td>brother</td>
<td>lmao</td>
</tr>
<tr>
<td>state</td>
<td>tomorrow</td>
<td>boobs</td>
<td>dr</td>
<td>sister</td>
<td>shit</td>
</tr>
<tr>
<td>texas</td>
<td>good</td>
<td>naked</td>
<td>dentist</td>
<td>uncle</td>
<td>ass</td>
</tr>
<tr>
<td>south</td>
<td>ill</td>
<td>wearing</td>
<td>tooth</td>
<td>cousin</td>
<td>bitch</td>
</tr>
</tbody>
</table>
Results
emotional

(a) Chain Frequency

PII & PEI

(b) Conversation Length
Results: Interpretation

- Emotional openness
  - When they are not very close, they express frequent encouragements, or polite reactions to baby or pets

Aww! What kind of puppy do you have?

I have a schnoodle named Harvey :) he's 10 months old now

Omg, schnoodles are so cute. I have 2 chocolate labs, they are puppies at heart

Aw those are great running dogs!!! My dog is a better walker :)
Results: Interpretation

- PII
  - When they meet new acquaintances, they use PII to introduce themselves.

I live between Camberwell, Peckham, Brixton & Dulwich.

I lived in Primros Hill *the Island* for fifteen years.
Results

Analyzing outliers: a dyad linked weakly but shows high self-disclosure

And now I can hear mum sobbing at dad and it’s too much and fuck I really just can’t take this.

tell me what happened D:

I just sort of snapped. It happens every now and again. She just always talks AT me and she WILL NOT SHUT UP. It’s overbearing.

ahh I understand that happens with my mom my grandmom my dad I do that :/ you feel so guilty :

ik the feeling. And that just makes it worse. That makes it hurt more.

yeah I hate the feeling :c it makes me sick XC
Conclusion

- Collected a large corpus of Twitter conversations
- Measured relationship strength by conversation frequency and conversation length
- Measured self-disclosure by
  - Negative, nonverbal, emotional openness
  - PII, PEI
  - Profanity
- Annotated PII and PEI using Mturk
- Confirmed hypothesis that stronger relationships show more self-disclosure behaviors
- Found some exceptions in emotional openness and PII
Thank you!
Any questions or comments?

JinYeong Bak
jy.bak@kaist.ac.kr
Department of Computer Science, KAIST
Backup slides
Results: Interpretation

- **Emotional openness**
  - When they are not very close, they express frequent greetings, encouragements, or polite reactions to baby or pets.

---

**Top 3 topics in weak relationships**

<table>
<thead>
<tr>
<th>Weak 1: greeting</th>
<th>Weak 2: encourage ment</th>
<th>Weak 3: baby/pets</th>
</tr>
</thead>
<tbody>
<tr>
<td>following</td>
<td>ill</td>
<td>love</td>
</tr>
<tr>
<td>thanks</td>
<td>sure</td>
<td>thanks</td>
</tr>
<tr>
<td>followers</td>
<td>soon</td>
<td>cute</td>
</tr>
<tr>
<td>welcome</td>
<td>better</td>
<td>aww</td>
</tr>
<tr>
<td>follow</td>
<td>want</td>
<td>pretty</td>
</tr>
</tbody>
</table>

---

- And now I can hear mum sobbing at dad and it's too much and fuck I really just can't take this.
- tell me what happened D:
- I just sort of snapped. It happens every now and again. She just always talks AT me and she WILL NOT SHUT UP. It's overbearing.
- ahh I understand that happens with my mom my grandmother my dad I do that:/ you feel so guilty :(
Future work

- Relationship strength
  - Network distance, community, relationship duration
- Emotion
  - Suggest new semi-supervised model
- PII and PEI
  - Time, place
  - Structure of questions and answers

And now I can hear mum sobbing at dad and it's too much and fuck I really just can't take this.

Tell me what happened D:

I just sort of snapped. It happens every now and again. She just always talks AT me and she WILL NOT SHUT UP. It's overbearing.
Method

- Self-Disclosure – Personal Information
  - Personally Identifiable Information (PII)
    - Information that can be used to uniquely identify a person
  - Personally Embarrassing Information (PEI)
    - Information that the damage that can be done by people that they know, the people for whom they might be embarrassed in front of

[boyd, danah. 2010]