Self-disclosure topic model for Twitter conversations

JinYeong Bak, Chin-Yew Lin, Alice Oh
jy.bak@kaist.ac.kr
Department of Computer Science, KAIST
Overview

- Self-disclosure (SD)
  - Definition from social psychology
  - Relations with social dynamics
  - Limitations in previous research

- Computational approaches for self-disclosure
  - Twitter conversation dataset
  - Self-disclosure topic model (SDTM)

- Self-disclosure & Social dynamics
  - Self-disclosure and conversation length over time
Self-disclosure (SD)
I want to be your friend

Where are you from?

I’m from Korea.
Self-disclosure

I want to be your friend

Where are you from?

I'm from Korea.

Can I just die? I hate everything

Don't die! I know how you feel
Self-disclosure

I want to be your friend

Where are you from?

I'm from Korea.

Can I just die? I hate everything

Don't die! I know how you feel

Today's my mother's birthday!

Tell her I said Happy Birthday
Self-disclosure

I want to be your friend.

Where are you from?

I'm from Korea.

Can I just die?
I hate everything!

Don't die! I know how you feel.

Today’s my mother’s birthday!

Tell her I said Happy Birthday.

My baby still has a fever!

Sending well wishes.
Self-disclosure: Definition

- The verbal expressions by which a person reveals aspects of self to others [Jourard1971b]
- Process of making the self known to others [Jourard&Lasakow1958]
Self-disclosure: Level

Self-disclosure level [Vondracek et al.1971, Barak et al.2007]

- General level (No disclosure)
- Medium level (Medium disclosure)
- High level (High disclosure)
Self-disclosure: G level

- General information and ideas
- No information about self or someone close to him

fabio capello is the manager are u sure its someone else whos playing lol

common you guys England manager is Roy Hodgson

noooooo we mean the manager before!
Self-disclosure: M level

- General information about self or someone close to him
- Personal events, age, occupation and family members

Today's my mother's birthday and she was extremely happy when I informed her I'm applying for Phoenix soon. Happy Birthday mom! :D

HAHA, nice! Tell her I said Happy Birthday and give her a kiss and hug for me! :3

That is a bit problematic. My mommy is not here lol
Self-disclosure: H level

- Sensitive information about self or someone close to him
- Problematic behaviors of self and family members
- Physical appearance, health, death, sexual topics

My mom has just been taken to the hospital by ambulance. Please pray for her. Thank you

Hugs. Glad your mom is doing better

Thanks, she is in hospital & is very disoriented.
Self-disclosure: Relations

Human relationship

- Degree of self-disclosure in a relationship depends on the strength of the relationship [Duck2007]
- Strategic self-disclosure can strengthen the relationship
Self-disclosure: Relations

Benefits

- Can get social support from others [Derlega et al. 1993]
- Can cope with stress [Derlega et al. 1993, Tamir and Mitchell 2012]

Examples

- My baby still has a fever argh!
- Sending well wishes ur way darling xo
- Aw thanks sweetness I appreciate it. Haven’t slept ugh
- My mom has just been taken to the hospital by ambulance. Please pray for her. Thank you
- Hugs. Glad your mom is doing better
- Thanks, she is in hospital & is very disoriented.
Limitations in Previous Works

- **Survey**
  - Asking questions to participants
  - Cons) Biased by participants memory

- **Hand coding**
  - Analyzing dataset by human
  - Cons) Cannot apply to large dataset
Limitations in Previous Works

- **Survey**
  - Asking questions to participants
  - Cons) Biased by participants memory

- **Hand coding**
  - Analyzing dataset by human
  - Cons) Cannot apply to large dataset

- **Lab environment**
  - Experiments held in lab or artificial environment
  - Cons) Not real/naturally occurring dataset
Research Questions

- How can we find self-disclosure in large & naturally occurring corpus automatically?
Research Questions

- How can we find self-disclosure in large & naturally occurring corpus automatically?
- What are relations between self-disclosure and social dynamics in large & naturally occurring corpus?
Twitter Conversations
Conversation in Twitter

Bak JinYeong @NoSyu

@smd4 Thanks to make excel2wiki.net It's really helpful for me!

Britney Spears @britneyspears

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! 🎶 Girl Gone Wild by Madonna — path.com/p/1zoiB

Madonna @Madonna

@britneyspears please come on stage and kiss me again. I miss you!!

Britney Spears @britneyspears

@MadonnaMDNAday Tempting...

Madonna @Madonna

@britneyspears Are you gonna make me work for this?

Britney Spears @britneyspears

@MadonnaMDNAday Why of course!

https://twitter.com/NoSyu
Users discuss several topics with others

Soccer

fabio capello is the manager are u sure its someone else whos playing lol

common you guys England manager is Roy Hodgson

nooooo we mean the manager before!

Politics

SENATOR STOP LYING! Falsely Claims Unemployment Is Higher Than When #Obama Took Office

It is higher, and that is the truth 22% in our county

Unemployment is not higher now then when Obama took office.
Conversation Topics

Users discuss several topics with others

Places

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

Family

- My baby still has a fever argh!
- sending well wishes ur way darling xo
- Aw thanks sweetness I appreciate it. Haven’t slept ugh
Twitter Conversations

- A Twitter conversation
  - 5 or more tweets
  - At least one reply by each user

Example of a Twitter conversation

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! 🎶 Girl Gone Wild by Madonna — path.com/p/1zoiB

@britneyspears please come on stage and kiss me again. I miss you!!

@MadonnaMDNAday Tempting...

@britneyspears Are you gonna make me work for this?

@MadonnaMDNAday Why of course!

https://twitter.com/britneyspears
Twitter Conversations

- A Twitter conversation
  - 5 or more tweets
  - At least one reply by each user

- Twitter conversation data
  - Aug 2007 to Jul 2013
  - 102K users
  - 2M conversations
  - 17M tweets

Example of a Twitter conversation

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! 🎶 Girl Gone Wild by Madonna — path.com/p/1zoIb

@britneyspears please come on stage and kiss me again. I miss you!!

@MadonnaMDNAday Tempting...

@britneyspears Are you gonna make me work for this?

@MadonnaMDNAday Why of course!

https://twitter.com/britneyspears
Self-disclosure Topic Model (SDTM)
Challenges for SD research

- Lack of ground-truth dataset of SD level
  - No tagged dataset for Twitter conversation
  - No accessible self-disclosure datasets
Challenges for SD research

- Lack of ground-truth dataset of SD level
  - No tagged dataset for Twitter conversation
  - No accessible self-disclosure datasets
- Lack of study about SD in computational linguistics
  - Definitions and examples in social psychology
  - Survey or hand-coding
  - Related word categories in LIWC [Houghton2012]
Ground-truth Dataset

Process

- Sample 301 random Twitter conversations
- Ask it to three judges
- Tag self-disclosure level
- Work on a web-based platform
Ground-truth Dataset

- **Process**
  - Sample 301 random Twitter conversations
  - Ask it to three judges
  - Tag self-disclosure level
  - Work on a web-based platform

- **Result**
  - Tagged G: 122, M: 147, H: 32 conversations
  - Fleiss kappa: 0.68
Assumptions: First person pronouns

First person pronouns are good indicators for self-disclosure

- Ex) ‘I’, ‘My’
- Used in previous research [Joinson et al. 2001, Barak et al. 2007]

I live between Camberwell, Peckham, Brixton & Dulwich.

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

Today's my mother's birthday and she was extremely happy when I informed her I'm applying for Phoenix soon.

Happy Birthday mom! :D

HAHA, nice! Tell her I said Happy Birthday and give her a kiss and hug for me! :3
Assumptions: First person pronouns

First person pronouns are good indicators for self-disclosure

- Ex) ‘I’, ‘My’
- Observed as highly discriminative features between G and M/H in annotated dataset

<table>
<thead>
<tr>
<th>Unigram</th>
<th>Bigram</th>
<th>Trigram</th>
</tr>
</thead>
<tbody>
<tr>
<td>my</td>
<td>I love</td>
<td>I have a</td>
</tr>
<tr>
<td>I</td>
<td>I was</td>
<td>is going to</td>
</tr>
<tr>
<td>I’m</td>
<td>I have</td>
<td>to go to</td>
</tr>
<tr>
<td>but</td>
<td>my dad</td>
<td>want to go</td>
</tr>
<tr>
<td>was</td>
<td>go to</td>
<td>and I was</td>
</tr>
<tr>
<td>I’ve</td>
<td>my mom</td>
<td>going to miss</td>
</tr>
</tbody>
</table>
Assumptions: Topics

M and H level have different topics

- [General vs Sensitive] information about self or intimate
## Assumptions: Topics

Self-disclosure related topics by LDA [Bak2012]

<table>
<thead>
<tr>
<th>Location</th>
<th>Time</th>
<th>Adult</th>
<th>Health</th>
<th>Family</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>
Assumptions: Topics

M and H level have different topics

- [General vs Sensitive] information about self or intimate
- Can be formalized as topics
  - Personally Identifiable Information
    - General information about self
    - Ex) name, location, email address, job, …
  - Secrets
    - Sensitive information about self
    - Ex) physical appearance, health, sexuality, death, …
Self-Disclosure Topic Model (SDTM)
Self-Disclosure Topic Model (SDTM)

- Classifying G and M/H level
  - Maximum entropy classifier
  - Observed first-person pronouns
Self-Disclosure Topic Model (SDTM)

- Classifying G and M/H level
  - Maximum entropy classifier
  - Observed first-person pronouns

- Classifying M and H level
  - Seed words for each level
  - Observed words

Graphical model of Self-Disclosure Topic Model
Self-Disclosure Topic Model (SDTM)

Rough figure of how to infer self-disclosure in SDTM

Tweet

Maximum Entropy Classifier

Topic Model

Topic Model with Seed Words

G level

M level

H level
Self-Disclosure Topic Model (SDTM)

Rough figure of how to infer self-disclosure in SDTM

Tweet

Maximum Entropy Classifier → Topic Model

G level

M level

H level

Topic Model with Seed Words
Self-Disclosure Topic Model (SDTM)

Rough figure of how to infer self-disclosure in SDTM

Tweet

Maximum Entropy Classifier

Topic Model

Topic Model with Seed Words

G level

M level

H level
Self-Disclosure Topic Model (SDTM)

Rough figure of how to infer self-disclosure in SDTM

Tweet

Maximum Entropy Classifier

Topic Model

G level

M level

H level

Topic Model with Seed Words
Self-Disclosure Topic Model (SDTM)

Rough figure of how to infer self-disclosure in SDTM
Self-Disclosure Topic Model (SDTM)

Rough figure of how to infer self-disclosure in SDTM

Tweet → Maximum Entropy Classifier → Topic Model → Topic Model with Seed Words

G level → M level → H level
Maximum Entropy Classifier

- Learned from annotated dataset
- Works better than others (C4.5, Naïve Bayes, SVM with linear kernel, polynomial kernel and radial basis)
- Used to identify aspect and opinions in topic model [Zhao2010]
Seed Words

Seed words are prior knowledge for each level

- G level
  - No seed words (symmetric prior)
- M level
  - Data-driven approach in Twitter conversation
- H level
  - Data-driven approach from external dataset
Seed Words

- M level
  - Data-driven approach
    - Use Twitter conversation dataset
    - Get frequently occurred trigram that begin with ‘I’ and ‘my’
Seed Words

- M level
  - Data-driven approach
    - Use Twitter conversation dataset
    - Get frequently occurred trigram that begin with ‘I’ and ‘my’
  - Example seed words

<table>
<thead>
<tr>
<th>Name</th>
<th>Birthday</th>
<th>Location</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>My name is</td>
<td>My birthday is</td>
<td>I live in</td>
<td>My job is</td>
</tr>
<tr>
<td>My last name</td>
<td>My birthday party</td>
<td>I lived in</td>
<td>My new job</td>
</tr>
<tr>
<td>My real name</td>
<td>My bday is</td>
<td>I live on</td>
<td>My high school</td>
</tr>
</tbody>
</table>
Seed Words

- H level
  - Data-driven approach
    - Use external dataset (*Six Billion Secrets*)
      - [http://www.sixbillionsecrets.com](http://www.sixbillionsecrets.com)
      - Users write and share his/her secrets
      - 26,523 posts
    - Extract high ranked word features

Example of secret posts in *Six Billion Secrets*
Seed Words

- **H level**
  - **Data-driven approach**
    - Use external dataset (*Six Billion Secrets*)
      - [http://www.sixbillionsecrets.com](http://www.sixbillionsecrets.com)
      - Users write and share his/her secrets
      - 26,523 posts
  - Extract high ranked word features

**Example seed words**

<table>
<thead>
<tr>
<th>Physical appearance</th>
<th>Health condition</th>
<th>Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>chubby</td>
<td>addicted</td>
<td>dead</td>
</tr>
<tr>
<td>fat</td>
<td>surgery</td>
<td>died</td>
</tr>
<tr>
<td>scar</td>
<td>syndrome</td>
<td>suicide</td>
</tr>
<tr>
<td>acne</td>
<td>disorder</td>
<td>funeral</td>
</tr>
</tbody>
</table>
Classifying Performance

- **Data**
  - Annotated Twitter conversation
  - 80/20 train/test randomly
Classifying Performance

- **Data**
  - Annotated Twitter conversation
  - 80/20 train/test randomly

- **Methods**
  - **BOW+**
    - Bag of Words + Bigrams + Trigrams features
  - **FirstP**
    - First-person pronouns features
  - **SEED**
    - Seed words and trigrams features
  - **FirstP+SEED**
    - FirstP and SEED feature
  - **SDTM**
    - Self-disclosure Topic Model
Classifying Performance

Accuracy

- BOW+: 54.1
- FirstP: 63.2
- SEED: 54.4
- FirstP+SEED: 60.4
- SDTM: 64.5
Classifying Performance

Accuracy

- BOW+: 54.1
- FirstP: 63.2
- SEED: 54.4
- FirstP+SEED: 60.4
- SDTM: 64.5
Classifying Performance

F1 Measure

- G F1
- M F1
- H F1
- Average F1

Bar chart showing F1 measure for BOW+, FirstP, SEED, FirstP+SEED, and SDTM.
Classifying Performance

F1 Measure

- G F1
- M F1
- H F1
- Average F1

Bar chart showing F1 Measure for different methods:
- BOW+
- FirstP
- SEED
- FirstP+SEED
- SDTM

Two methods highlighted with blue circles.
Classifying Performance

F1 Measure

<table>
<thead>
<tr>
<th>Method</th>
<th>G F1</th>
<th>M F1</th>
<th>H F1</th>
<th>Average F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOW+</td>
<td>0.5</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>FirstP</td>
<td>0.6</td>
<td>0.7</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>SEED</td>
<td>0.5</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>FirstP+SEED</td>
<td>0.6</td>
<td>0.7</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>SDTM</td>
<td>0.5</td>
<td>0.6</td>
<td>0.4</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Classifying Performance

F1 Measure

- G F1
- M F1
- H F1
- Average F1

Bar chart showing performance metrics for different methods: BOW+, FirstP, SEED, FirstP+SEED, SDTM.
Self-disclosure & Social dynamics
Research Questions

Q1) Does high self-disclosure lead to longer conversations?

Can I just die?
I hate everything

Don't die!
I know how you feel

Here is gift for you

I like you too!

Thanks!

You're my best friend!
Research Questions

Q2) Is there difference in conversation length patterns over time depending on overall self-disclosure level?
## Results

High ranked topics in each level (G, M, H levels)

Shown by high probability words in each topic

<table>
<thead>
<tr>
<th>G1</th>
<th>G2</th>
<th>M1</th>
<th>M2</th>
<th>H1</th>
<th>H2</th>
</tr>
</thead>
<tbody>
<tr>
<td>obama</td>
<td>league</td>
<td>send</td>
<td>going</td>
<td>better</td>
<td>ass</td>
</tr>
<tr>
<td>he’s</td>
<td>win</td>
<td>email</td>
<td>party</td>
<td>sick</td>
<td>bitch</td>
</tr>
<tr>
<td>romney</td>
<td>game</td>
<td>i’ll</td>
<td>weekend</td>
<td>feel</td>
<td>fuck</td>
</tr>
<tr>
<td>vote</td>
<td>season</td>
<td>sent</td>
<td>day</td>
<td>throat</td>
<td>yo</td>
</tr>
<tr>
<td>right</td>
<td>team</td>
<td>dm</td>
<td>night</td>
<td>cold</td>
<td>shit</td>
</tr>
<tr>
<td>president</td>
<td>cup</td>
<td>address</td>
<td>dinner</td>
<td>hope</td>
<td>fucking</td>
</tr>
</tbody>
</table>
Results

Q1) Does high self-disclosure lead to longer conversations?
Ans) Positive relations between initial SD level and changes CL

![Graph showing the relationship between initial SD level and changes in conversation length]
Results

Q2) Is there difference in CL patterns over time by overall SD level?

Ans) ‘high’ and ‘mid’ groups increase CL over time, not ‘low’

‘high’ groups talk more in a conversation than ‘mid’ & ‘low’ groups
Contributions

- Made ground-truth Twitter conversation dataset for SD level
  - Made first annotated Twitter conversations for SD level
  - Share it with researchers

- Suggested novel method for identifying SD level (SDTM)
  - Our assumptions are reasonable and verified by experiments
  - SDTM performs better than others

- Showed relations between SD & social dynamics
  - Strategic self-disclosure can strengthen the relationship
    supported by Twitter conversation dataset and SDTM
Future Work

Self-disclosure for a user general messages

- Self-disclosure is related with
  - Loneliness [Al-Saggaf.2014]
  - Online social network usage [Trepte2013]

- We can predict user’s
  - Loneliness and give a social support
  - Usage patterns and give a feedback
Reference

Thank you!
Any questions or comments?

JinYeong Bak
jy.bak@kaist.ac.kr
Department of Computer Science, KAIST