Discovering Dedicators with Topic-based Semantic Social Networks

Jiyeon Jang (IR&NLP LAB @ KAIST)
with Sung-Hyon Myaeng (KAIST)

2013.07.09.
Motivation

• In order to spread information successfully, some extraordinary people are needed [Gladwell, 2002].

Information specialist

Maven ≈ expert

Salesman ➔ Dedicator!

Persuader

Connector ≈ influencer

People who know a large number of people

Dedication: the extent to which a user has dedicated to transmit information in selected topic areas to the people in their egocentric social networks
Influencer vs. Deducator

Influencer vs. Deducator

Deducators are right people to deliver certain information to their friends directly! (Real Contributor)

→ Exposure

Do not know whether the green users read it or not

→ getting people engaged in an issue

Via “conversation”
Topic-dependent Dedication

- Our approach is to analyze Twitter conversations
  - The actual flow of information among people

- Online Conversations promote two-way information exchange of certain issues

- Conversations usually have one or more topics
  → We look for topic dependent dedicators

Jiyeon Jang (IR&NLP LAB @ KAIST)
Research Goal

• How to measure dedication of a user for certain topic areas?
• How does the dedicator differ from the influencer?
• What are the characteristics of the dedicators?
Twitter Conversation Dataset

• Dataset Collection Process

  Collect “public timeline” of Twitter at 30s interval
  09/09/2011 ~ 10/04/2011

  Keep **users** having “@” tweet, 3,200 tweets in total

  Recover **all the conversations** of each user

  Filter out users having < 400 conversations

  Remove stopwords, unusual meaningless words such as “aaaa”, “mossst”

<table>
<thead>
<tr>
<th>Volume of Dataset</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#Users</td>
<td>1,550</td>
</tr>
<tr>
<td>#Conv</td>
<td>1M</td>
</tr>
<tr>
<td>#Tweets</td>
<td>6M</td>
</tr>
</tbody>
</table>
Topic-based Semantic Social Network

• In order to measure topic-dependent dedication

• Ego-centric topic-based semantic social networks [Jang et al. 2012]
  • Each relationship is represented as a topic distribution

  ![Diagram showing topic distribution between UserA and UserB]

  • Topic Diversity: coverage of topics
  • Topic Purity: topical focus
Use of Common Global Topics

- Topic spectrums are different from user to user
- Need for common topics to compare dedication levels
- Keep individual topics while mapping them to global topics

- Our approach

1,550 users * 50 topics per user
Use of Common Global Topics

- Topic spectrums are different from user to user
- Need for a common topics to compare dedication levels
- Keep individual topics while mapping them to global topics

An example of a global topic and associated individual topics

<table>
<thead>
<tr>
<th></th>
<th>Top-ranked words representing a topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Topic</td>
<td>{song, listen, music, sing, album, ... }</td>
</tr>
<tr>
<td>User1’s topic</td>
<td>{like, just, song, new, album, ... }</td>
</tr>
<tr>
<td>User2’s topic</td>
<td>{song, love, like, album, listen, ... }</td>
</tr>
<tr>
<td>User3’s topic</td>
<td>{like, listen, music, song, panic, ... }</td>
</tr>
<tr>
<td>User4’s topic</td>
<td>{kim, love, lil, album, plai, ... }</td>
</tr>
<tr>
<td>User5’s topic</td>
<td>{queen, selena, listen, album, think, ... }</td>
</tr>
</tbody>
</table>
Dedication Factors

• To measure a user’s dedication level for a topic
• Volume
  # of conversations, # of conversation partners
• Engagement
  # of conversations per partner, # of tweets in a conversation, inverse of time lag between conversations
  Inverse of time lag between tweets in a conversation
• Personal Tendency
  topic diversity, topic purity
• Topic Weight
  topic probability
Absolute vs. Relative Level Dedication

- Need to consider both absolute (community level) and relative (individual level) dedication

Comparing with others

We combine absolute level dedication and relative level dedication when measuring dedication level of a user!
Analysis
Correlation Analysis

• Uniqueness of Factors

<table>
<thead>
<tr>
<th>Factors Compared</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume vs. engagement</td>
<td>0.182</td>
</tr>
<tr>
<td>volume vs. tendency</td>
<td>0.143</td>
</tr>
<tr>
<td>engagement vs. tendency</td>
<td>-0.042</td>
</tr>
</tbody>
</table>

p<0.01

• The three factors capture different aspect of dedication

<table>
<thead>
<tr>
<th>Factors Compared</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume vs. topic weight</td>
<td>-0.006</td>
</tr>
<tr>
<td>engagement vs. topic weight</td>
<td>0.116</td>
</tr>
<tr>
<td>tendency vs. topic weight</td>
<td>0.546</td>
</tr>
</tbody>
</table>

p<0.01
Characteristics of Dedicators

- Comparison between Dedications and Influences
  - Influence: In-degree influence (# of follower links)

<table>
<thead>
<tr>
<th>Influence vs. dedication</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In-degree vs.</strong> Dedication (absolute &amp; relative levels)</td>
<td>-0.030</td>
</tr>
<tr>
<td><strong>In-degree vs.</strong> Dedication (absolute level)</td>
<td>-0.005</td>
</tr>
<tr>
<td><strong>In-degree vs.</strong> Dedication (relative level)</td>
<td>-0.118</td>
</tr>
</tbody>
</table>

*p<0.01

No Correlation!

- In-degree influence & dedication measure different aspects of an individual
Characteristics of Dedicators

• Comparison between Dedicators and Influencers
  • Example
    • Topic about music ({song, listen, music, sing, album, ...})
    • 77 users have communicated on this topic

No overlap between them

Top-5 dedicators

Top-5 indegree influencers

Avg. Indegree (# follower)
Volume (Absolute)
Volume (Relative)
Engagement (Absolute)
Engagement (Relative)
Personal Tendency (Absolute)
Personal Tendency (Relative)

0.77

0.75
Characteristics of Dedicators

• Role of Factors Contributing to Dedication

In-degree Influence makes the least contribution
Summary

• Proposed the concept of dedicators in social networks referring to the theory of “The Law of the Few”

• Findings
  • Users with high in-degree influence tend not to be strong dedicators
  • Top dedicators tend to have richer conversations, taking advantages of smaller and manageable social networks

• Application
  • Identifying real contributors in information dissemination in marketing
Thank you!

Any Questions?