MIB605- Lecture 9
Social Media Research

Prof. Cui

Agenda

▶ Managing Social Media Communities (MS Ch7 and Ch8)
  ▶ Quality of Opinion Community
  ▶ Minimizing Social Dynamics

▶ Monitoring Social Media Metrics (TS Ch10)

▶ Social Media Research (TS Ch9)
  ▶ Qualitative vs. quantitative social media research
  ▶ Sentiment and Content Analysis

▶ Hamilton, Schlosser, and Chen’s 2017 paper
Objetives

• **Why** to monitor?
  - Learn from the community
  - Evaluate effectiveness of social media campaign
  - Justify the value of investing in social media activities

• **What** to measure?

• **How** to measure?

• **Data mining, mining intelligence, big data analytics**

Firms’ management of social interaction

► Firms’ management of social interaction (Godes et al. 2005); from passive to aggressive:
  - The firm as observer
    - Collects social interaction information to learn about its ecosystem
  - The firm as moderator
    - Fosters social interactions
  - The firm as mediator
    - Actively manages social interactions
  - The firm as participant
    - Plays a role in the social interactions

► No matter what role we are playing, we need to “evaluate” our performance. Thus, how should we evaluate?
Should company trust opinions from social media?

Yes, and No!

**Strength:** Large source of information available on the Internet

**Problems:**
- Whether posters have the true experience (e.g., fake review?)
- Whether posters have the knowledge (e.g., review capability)
- Whether posters’ opinion represent the whole population (e.g., selection bias, motivation, participation inequality)
- Whether posters are expressing their own opinions (e.g., social influence)
  - Trolls, bots, incentivizers, freelancers, and the invisible hand of big brother!

**Objectives:** To make online opinions more transparent, trustworthy, representative! To provide more intelligence to firms.

---

Evaluation based on what consumers say, but….

- Everyone can post online, and their opinions may be biased, or they can’t represent the opinions from our “target customers”. => *The democratic nature allows everyone to express opinions.* (given being anonymous)

- Still, these opinions affect the community.

- Then, early discussion can direct the general public (e.g., bandwagon effect, social dynamics).
Quality of Opinion Community

A. Increasing the validity of reviews- readers “believe” the reviews!

1. Accountability vs. anonymity (via verified purchase, disclosure of identity)
2. Reputation Management (by rewarding top reviewers with status, votes of helpfulness)
3. Verify and certify reviews
4. Negative reviews (suppression or proactive responding)

B. Minimizing Social Dynamics- independent reviews are preferred!

1. Encourage a variety of voices
2. Minimize expert effects
3. Impact of social dynamics

A. Increasing the validity of reviews

Increasing the validity of reviews- readers “believe” the reviews!

1. Accountability (identity) vs. anonymity
Quantitative evaluation scale may be the same, but anonymity can degrade the quality of textual reviews. (more extreme, less dispersion)

However, for personnel evaluations, anonymous opinions tend to be more negative and critical.

<Discussion> So, which approach is better? (only members, verified purchases)
A. Increasing the validity of reviews

2. Reputation Management

Idea of “pseudonymity”- real identity is hidden to the public, but is known to the website.

Reputation management systems allow users to “grade the credibility of others.” (e.g., eBay, reviewing a seller/buyer; Amazon)

Votes of helpful/usefulness of reviews

---

A. Increasing the validity of reviews

3. Verify and certify reviews

People may post without real experience/knowledge. Amazon would verify if one is “verified purchase”- meaning that the person really purchased the item from Amazon.

4. Negative reviews: true or not (cannot delete by definition, but,..)

Build a filtered community there negative reviews are systematically deleted: illegal and unethical! (aside from censorship!)

For consumers, lack of opinion variation may suppress further discussion. For firms, biased and incomplete view is collected.
B. Minimizing Social Dynamics

1. Encourage a variety of voices
   Self-selection effects are most pronounced when “moderate opinions” are held by the majority but the comments being posted come from a vocal minority whose views may differ substantially from those of an average individual.

The silent majority vs. the loud minority

Solutions?
Send follow-up email to buyers.
Provide incentives for customers to post an online review.

2. Minimize expert effects (砖家)

Experts tend to be more critical and posting negative reviews.
-> direct them to focus on “being helpful” rather than “signaling their expertise”.

Example: Amazon’s questions about “Was this review helpful to you?”
B. Minimizing Social Dynamics

3. Impact of social dynamics
Website A: people agree on the product quality
Website B: more variation in opinions
What happened for the future review?
Consumers post more varied opinions on website B than website A, even the “average opinions” are the same.
Sales, on the other hand, are greater on website B.

Summary:

- Manage the opinion environment to minimize the biases.
  - Non-anonymous posting / ID linked to Facebook or your social network
  - Require more profile info.
- Encourage participation from a broader population.
- Moderate conversations to minimize unwanted dynamics and social influences.
- Design incentives to create a community that fosters a healthy dialog.
Measurement

- Marketing Metric - A measure of the **quantitative value** or trend of a marketing action or result
- Social Media Metric - A measure of the **quantitative value** or trend of a **social marketing** action or result

What Matters is Measured

What to measure? *Mimic online advertising:*

- **Reach** (# of people exposed to the message)
- **Frequency** (average # of times someone is exposed)
- **Relative pull** (how different creative executions generate a response)
- **Sales conversion** (# of people who buy the product)
- **Viewthroughs** (# of people who later visit the brand’s websites)

However,

1) These may not relate to the goal of social media;
2) All metrics are **quantitative**, rather than **qualitative**.
What to measure? Commonly-used Social Media Metrics

1. Buzz volume
   • # of posts, comments, retweets, etc.
   • Frequency, momentum, recency, seasonality.

2. Asset popularity, virality
   Sharing, viewing, bookmarking, downloads, installs, and embedding

3. Media mentions (earned media)

4. Brand liking
   • Fans, followers, friends
   • Growth in fans, followers, friends
   • Likes, favorites, ratings, links back

What to measure?

5. Reach and second degree reach (influence impressions from others)
   • Readers, viewers
   • Subscriptions
   • Mentions, links

6. Engagement
   • Comment volume
   • Uploads, contest participation
   • Registration
   • Time-spent
   • Subscription (RSS, podcasts, video series, channel)

7. Sentiment (content analysis)

8. Website effectiveness (traffic, clicks, conversions, viewthroughs)
Engagement (example of \textit{qualitative measure})

\textbf{The engagement level:}
Saw -> Saved -> Rated -> Repeated -> Commented -> Clicked ->
Interacted -> Purchased -> Recommended

Not all “data” are relevant, and we should focus on “key performance
indicators (KPIs).
Example: If your social media marketing goal is to get people comment, then
focus on the measurement of comment.

1. \textit{Measurements within a defined context of metrics.}
2. \textit{Measurements require context to provide useful feedback.}
3. \textit{Metrics that are tied to objectives are key performance indicators.}
4. \textit{Objectives must be well-defined before we can identify key performance indicators.}

\textbf{Social Media Marketing Matrix}

\textit{Activity} metrics: measure the actions the firm takes relative to social media.

\textit{Interaction} metrics: how that target market engages with the social media
platforms and activities.

\textit{Return} metrics: focus on the outcomes (e.g., financial or otherwise) that
directly or indirectly support the success of the social media activity.

\textit{Social Media Return on Investment} (SMROI)
HOW: The evaluation and measurement process (DATA)

Define: define the results that the program is designed to promote
Assess: assess the costs of the program and the potential value of the results
Track: track the actual results and link those results to the program
Adjust: adjust the program based on results to optimize future outcomes.

(1) Define

Objective-> SMART

Specific
Measurable
Appropriate
Realistic
Time-oriented

Ex1:
“We will tell everyone we care about our new Facebook page and see if they like it so much as they’ll buy more of our product.”

Ex2:
“We will promote our new Facebook page in print ads, we will place in the June issues of Rolling Stone, Sports Illustrated, and Maxim. On July 15 we will count the number of Facebook users who “like” our brand and compare sales to the same period last year.”
(2) **Assess**

Cost-benefit analysis

Consider the calculation of return on investment (ROI)

---

(3) **Track**

a. Forward tracking: develop a tracking mechanism prior to launching the campaign.

b. Coincident tracking: no unique tracking mechanism is needed, and is effective in knowing the immediate effect of a campaign.

c. Reverse tracking: conducted after an activity or campaign has concluded. (E.g., do a survey to review the performance)
(4) Adjust

Results as planned? Profitable?

Any deviated outcomes should be interpreted, and adjust for future plan.

Social Media Research

1) Secondary vs. primary research

2) Quantitative vs. qualitative research

3) Content analysis
Primary Data Versus Secondary Data

- **Primary Data**
  - Newly collected data
  - Collected for the research problem by the researchers
  - Comes later than secondary data!
  - Method: Lab experiment, filed experiment, surveys, observation, focus groups, etc.

  *Primary data does not exist until you conduct a research project to collect it!*

---

Secondary Data

- **Data not gathered for the immediate study; the data already exists, having been collected for some other purpose.**
  - Previously collected
  - Might have been collected for a different purpose
  - Internal and external databases are both good sources of this data
  - When we use, no direct access to subjects
Comparison

<table>
<thead>
<tr>
<th></th>
<th>Primary Data</th>
<th>Secondary Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection Purpose</td>
<td>For the problem at hand</td>
<td>For other problems</td>
</tr>
<tr>
<td>Collection Process</td>
<td>Very involved</td>
<td>Rapid and easy</td>
</tr>
<tr>
<td>Collection Cost</td>
<td>High</td>
<td>Relatively low</td>
</tr>
<tr>
<td>Collection Time</td>
<td>Long</td>
<td>Short</td>
</tr>
</tbody>
</table>

Qualitative Social Media Research

A. Observational Research
   Study “what”, “when”, “how”, and “where” do people talk online.

   - Advantages
     - Understand consumers without “asking” them.
     - Alternate way to collect primary data

   - Disadvantages
     - Public behavior only
     - Time-consuming
     - Some “items” are more likely to discuss offline than online (see handout)
What Can Be Observed?

<table>
<thead>
<tr>
<th>Phenomena</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human behavior</td>
<td>Shoppers browsing pattern on a website</td>
</tr>
<tr>
<td>Expressive behavior</td>
<td>In a review, post: tone of voice, and the selection of words, photos, etc.</td>
</tr>
<tr>
<td>Temporal pattern</td>
<td>How often does one post? The gap between each post</td>
</tr>
</tbody>
</table>

Qualitative Social Media Research

B. Ethnographic Research

Study target people behavior in a natural and real-world. => popular in studying online communities.

How conversation evolves and size of community grows.

- Netnography: a methodology that adapts ethnographic research techniques to study the communities that emerge through *computer-mediated communications*.
- Chat room, online forums, message broad.
- Web crawling technique
Quantitative Social Media Research

A. Monitoring and Tracking

B. Sentiment Analysis (opinion mining)- bottom-up approach to study how people think or feel about an object.

C. Content Analysis- top-down approach to examine a proposed theory or evidence.

Sentiment Analysis, text mining!

A. Web scraping to get the “data” and filter the data.

B. Extract sentiment: using sentiment indicators by establishing (or using) a sentiment dictionary

C. Aggregate raw sentiment data into a “summary”.

Example: Pennebaker Linguistic Inquiry and Word Count (http://www.liwc.net/tryonline.php) with established sentiment dictionary Challenges with Chinese languages!
Content Analysis

Used to identify the presence of concepts and themes within qualitative data sets, uses a “top-down” approach that applies theory or empirical evidence to the coding process.

Example: A study propose that “female”, compared to “male”, is more likely to mention “other” consumers in a review, because they tend to care more about other people.

Coding: count the frequency of thoughts/wordings about other consumers.

Content Analysis- Exercise from Tripadvisor

Please read the “review” carefully, and code whether the poster mention the following attributes:

- Price
- Location
- Service
- Transportation
- Safety

If the hotel attribute is mentioned, coded it as “1”, otherwise “-1”.
Content Analysis

Coding categories for content analysis

<table>
<thead>
<tr>
<th>Type of Code</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context codes</td>
<td>Provide information on the source of the comment</td>
</tr>
<tr>
<td>Respondent perspective codes</td>
<td>Captures the general viewpoint revealed in the comment</td>
</tr>
<tr>
<td>Process codes</td>
<td>Indicate when over the course of the campaign a comment occurred</td>
</tr>
<tr>
<td>Relationship codes</td>
<td>Indicate alliances within social communities</td>
</tr>
<tr>
<td>Event codes</td>
<td>Indicate unique issues in the data</td>
</tr>
<tr>
<td>Activity codes</td>
<td>Identify comments that require response</td>
</tr>
</tbody>
</table>

Summary

Social intelligence relies on the understanding of what consumers are talking on social media.

We should systematically “manage” the platforms to get quality data.

The monitoring of social activities depends on our objective, and the cost-benefit analysis.

Different methodologies should be applied to analyze the data.
Seeing Is Believing: The Effects of Profile Pictures in Online Platforms

Yuho Chung, Lingnan University, Hong Kong, yuhochung@ln.edu.hk
Geng Cui, Lingnan University, Hong Kong, gcui@ln.edu.hk
Ling Peng, Lingnan University, Hong Kong, lingpeng@ln.edu.hk
Source credibility

- Processing motivation
- Credible sources are more trustworthy, persuasive
- Expertise, and motives
- Social cues (nonverbal)
- Aesthetics.

Source attractiveness

- Familiarity, similarity, and likability
- Identification
- Competent, acceptable, persuasive
- Can be distractive
- Advertising, celebrity endorsement
Impression formation online

- Disclosure the identity of reviewers and sellers greatly affects their credibility and persuasiveness.
- Ambiguity reduction and positivity of impression.
- Presence of profile pictures.

Research framework
Research framework

Pictorial Dye

- Technical aspect of pictorial identity information, the medium that facilitates intimacy and immediacy between information creators and receivers.
- Poor quality profile pictures that are low resolution, dark, small, blurry, and/or noisy may weaken the effects of visual stimuli and thus hinder interpersonal communication, leading to weaker social presence and normative influences.

(Dark & blurry image)  (Bright & sharp image)
Pictorial Proximity

- Psychological aspect of pictorial identity information, the perceptual distance of a pictorial object in a receiver’s eye.

\[
\text{Facial prominence ratio} = \frac{\text{Area of a face}}{\text{Area of a picture}}
\]

![ FP ratio: 2% ]

![ FP ratio: 16% ]

![ FP ratio: 90% ]

Pictorial Valence

- Physical aspect of pictorial identity information, the facial attractiveness of a pictorial object in a receiver’s eye

- Use a five-step process to determine pictorial valence:
  - (1) collect sufficient pictorial identity information from digital platforms;
  - (2) recruit raters to score facial attractiveness in a sample of profile images;
  - (3) use image processing techniques to retrieve key pictorial features;
  - (4) develop a machine learning algorithm to learn the relationship between geometric facial features and facial attractiveness; and
  - (5) score the facial valence of all profile images using the most accurate machine learning models
### Pictorial Valence

<table>
<thead>
<tr>
<th>Attractiveness score between 2.0 and 2.5</th>
<th>Attractiveness score between 3.0 and 3.5</th>
<th>Attractiveness score between 4.0 and 4.5</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Images" /></td>
<td><img src="image2.png" alt="Images" /></td>
<td><img src="image3.png" alt="Images" /></td>
</tr>
<tr>
<td><img src="image4.png" alt="Images" /></td>
<td><img src="image5.png" alt="Images" /></td>
<td><img src="image6.png" alt="Images" /></td>
</tr>
<tr>
<td><img src="image7.png" alt="Images" /></td>
<td><img src="image8.png" alt="Images" /></td>
<td><img src="image9.png" alt="Images" /></td>
</tr>
<tr>
<td><img src="image10.png" alt="Images" /></td>
<td><img src="image11.png" alt="Images" /></td>
<td><img src="image12.png" alt="Images" /></td>
</tr>
</tbody>
</table>

### Pictorial Tie

- Social aspect of pictorial identity information, the presence of social relationship in the pictorial information

- Single portrait
- Two-person portrait
- Group portrait
Methods

• Online reviews: Yelp
• Sharing economy: Airbnb
• E-commerce: 5miles
• Overall 600,000 profile pictures

Conclusion

• pictorial identity information serves as a predictor of social outcomes and financial performance, beyond those attributable to the characteristics of primary information (i.e., message content) or identity-descriptive source information.

• demonstrate strong evidence of the influence of pictorial identity information on social behavior, such as perceived review usefulness and financial performance (e.g., occupancy rate).

• show that pictorial identity information can improve sales performance and accelerate the purchase decision process by reducing buyers’ uncertainty on e-commerce platforms.
Break

- Followed by leading discussion group 6