Charting the Landscape of Chinese Social Media: What We Know and What We Don’t Know from Existing Research

Australian National University, August 1, 2014

Jonathan J. H. Zhu
with T. Q. Peng, H. Liang, J. Qin, Z. Z. Wang, H. X. Chen, & B. L. Li
Mapping Landscape of Chinese Social Media

Missing pieces:
• Knowledge sharing
• Games
• ?

Mapping Research on Chinese Social Media?
Diverse Research Outlets

- Official statistical reports:
  - e.g., CNNIC biannual surveys
- Listed-firm financial reports:
  - e.g., Sina, Tencent, etc.
- Web traffic monitors:
  - e.g., Alexa, Dratio (writing.com), etc.
- Marketing/consulting research reports:
  - e.g., McKinsey, iResearch, etc.
- Academic research papers:
  - English language; Chinese language
Accurate Data vs. Insightful Analysis

Accuracy

Insightfulness

Financial Reports

Consulting Analysis

Academic Papers
Revenue Model of Top 4 Social Media, 2011

Sources: the 2011 financial reports of Tencent, Easenet, Sina, and Sohu
CNNIC: The 34th Internet User Survey Report (July 2014)

By June 2014:

- Total users = 632M
- Penetration rate = 47%
- Mobile users = 83%
McKinsey Global Institute Report (July 2014)
China’s digital transformation: The Internet’s impact on productivity and growth

Acknowledgement:
“... This independent MGI initiative drew on data and expertise from the Alibaba Group Research Center and Baidu Development Research Center. Part of our analysis was made possible through this collaboration...”

Cover-page data:
- 632 million Chinese Internet users in 2014
- 700 million Active smart devices
- ~US$300 billion 2013 e-tailing sales
- 4.4% Size of China’s Internet economy as a share of 2013 GDP—higher than the United States or Germany
“Readers seeking predictions about how all of these changes will affect political outcomes in China over the next few decades are likely to be frustrated. This is not intended as a predictive book, and Mr. Yang wisely makes no attempt to forecast future events based on current trends. But for an in-depth understanding of how life has changed for China’s 300 million plus Internet users, and how these personal transformations have in turn affected contemporary Chinese society and culture, Mr. Yang’s work is essential reading.”
Growth of SSCI Publications on Social Media

- y-axis: N of Total Articles per Year
- x-axis: Years from 2006 to 2013

Lines:
- Blue: Total
- Red: China
Disciplines Involved in Social Media Research

- Other
- Humanities
- Communication
- Health & Environment
- Education
- Social Sciences
- Business & Economics
- Science & Technology
- Computer & Information

Total (N=4600+)

China (N=280)
The 5W Model for Assessing Research on Chinese Social Media

Harold Lasswell (1948): Communication is a process of who says what to whom through which channel with what effects.

Communicator $\rightarrow$ Content $\rightarrow$ Medium $\rightarrow$ Audience $\rightarrow$ Effects
W1: Who (Communicator)

What we are told:
- media professionals
- elite opinion leaders
- government mouthpieces
- rumor mills (“50-cent Party”, “Water Army”, Spammers, etc.)

What we aren’t told:
- grassroots opinion leaders
- e-marketers
- Foreign individuals/entities
Communicators on Early Weibo (N=10)

Government Presence on Social Media

- **Unique features:**
  - Centralized review process
  - Widespread “Water Army”
  - Adaption of international practices

- **Fundamental problem:**
  - Open-networked society vs. closed-centralized government

Growth of Weibo Accounts by Police Bureaus

Predictors of Use of Weibo by Police Bureaus

**Significant predictors:**
- Government size
- Population size
- Adoption of neighbors
- Adoption of upper-level

**Non-significant predictors:**
- Revenue
- GRP
- Openness
- Safety level
- E-government efforts
- User size

W2: Says What (Content)

What we are told:
- multiple duplicates
- contentious
- censorship

What we aren’t told:
- benchmark categorization
- non-political content (e.g., entertainment)
- forced vs. self-censorship
Categories of Early Weibo Posts (N=438)

- News: 31%
- Diaries: 28%
- Jokes: 19%
- Quotes: 10%
- Promotions: 6%
- Others: 4%
- Comments: 3%

Who Say What on Early Weibo

## Sina Weibo vs. Twitter

<table>
<thead>
<tr>
<th></th>
<th>Weibo</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trending Topics</td>
<td>Trivial Jokes</td>
<td>News events around the world</td>
</tr>
<tr>
<td>Trend Setters</td>
<td>Unverified Accounts</td>
<td>Leading News Media</td>
</tr>
<tr>
<td>Trendy Topics Retweeted</td>
<td>62%</td>
<td>31%</td>
</tr>
<tr>
<td>Presentation content</td>
<td>Images, Videos, Links</td>
<td>Text</td>
</tr>
</tbody>
</table>

What Might Cause the Differences?

- The higher rate of retweeting is likely to be created by spammers, i.e., the accounts that were set up to retweet (promote) targeted users, which also promote the relevant topics to be trending.
- About 1% of the users are suspected spammers, who contribute 1/3 of the total posts or half of the retweeted posts.
- If the spammers are removed, the evolution patterns of trending topics on Weibo becomes similar to that on Twitter.

Detecting Spammers by User-Retweet Ratio

N of Users Retweeted by the Account

% of Sample

Active Account
Inactive Account
Chinese Censorship

King et al. (2013):

- The largest selective suppression of human expression in history:
  - implemented manually (within a few hours of posting)
  - by 200,000 workers
  - located in government and inside social media firms

- A huge censorship organization:
  - (obviously) designed to suppress information
  - (paradoxically) very revealing about the goals, intentions, and actions of the Chinese leadership
What’s Censored

Source: King et al. (2013).
W3: Which Channel (Medium)

What we are told:
- Weibo concentrated
- compared with Twitter
- compared with old media

What we aren’t told:
- other social media
  - IM (WeChat)
  - BBS (Tianya)
  - SNS (QQ)
  - blogs, dating, gaming, etc.
- competition/cooperation among social media
Blogs vs. Microblogs

Method:
- 300 Sina blog posts on traveling to Hong Kong
- 300 Sina Weibo posts on the same topic
- Manual content analysis of the frequency, direction, and intensity of keywords, themes, visuals, etc.

Findings:
- Blog posts largely to recount past experience
- Weibo posts mainly to express desire to visit in advance
- No formal comparison made as the two samples treated as replicates rather than contrasts

Source: Tse & Zhang (2012). Analysis of Blogs and Microblogs
Competitor or Partners

PPM Model:

- Push Forces (satisfaction)
- Pull Forces (alternative attraction, etc.)
- Mooring Forces (perceived popularity, etc.)

Method:

- Drew a random sample of Sina users
- Mined their blog/Weibo posts to measure behavior
- Invited them to answer an online questionnaire to measure motivation
- Found most users to co-use both simultaneously

Source: Lu (2013). The sustainability of UGC.
W4: To Whom (Audience)

What we are told:
- massive size
- skewed participation
- polarized views

What we aren’t told:
- user authenticity
- passive audience
- inactive audience
Massive Size of Social Media Users

Sources: CNNIC survey report; Sina/Tencent annual reports
Changes in Popularity of Chinese Social Media

Source: CNNIC Annual Survey Reports (www.cnnic.cn)
## Distinctions between Communicator and Audience

<table>
<thead>
<tr>
<th>Participation Duration</th>
<th>Content Creation</th>
<th>Content Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td>I. Active Communicator</td>
<td>III. Active Audience</td>
</tr>
<tr>
<td>Interrupted</td>
<td>II. Inactive Communicator (traceable)</td>
<td>IV. Inactive Audience (completely ignored)</td>
</tr>
</tbody>
</table>

Political Preference of Dominant Groups

A 3-step Flow Model based on Li Chengpeng Debate

Source: Chen (2013). Opinion Leaders: The Driving Force of Political Discussion in Social Media
Geographic Disparity of Social Media Usage

Region Distribution of Social Sharing Users in 2013

W5: With What Effects

What we are told:
- registration = exposure
- exposure = conversion
- expression = participation

What we aren’t told:
- definition of effects
- links to content
- links to offline life
- effects on individuals vs. society
Expression vs. Participation

Arguments: Weibo helps rural migration workers to
- mobilize to improve their life in cities;
- release emotional dissatisfactions with reality;
- influenced by journalists, scholars, and officials (unreciprocal weak ties?).

Method:
- Searched “new generation of rural workers” (新生代农民工) to obtain 4,000+ posts;
- Content analyzed the themes of the posts;
- Results considered to support/derive the arguments.

Source: Zhang (2013). Social inclusion or exclusion?
Evidence (?) for Arguments 1-2

Source: Zhang (2013). Social inclusion or exclusion?
Agenda-setting Effects based on N of Posts

Hypotheses:
- “Safety valve” effects: bloggers express discontent on issues initiated by the media;
- “Pressure cooker” effects: bloggers lead issues to increase social tensions.

Method:
- Randomly searched through Google 2,000+ blog posts and 4,000+ newspaper stories;
- Computerized content analysis of issue themes;
- Time series analysis to determine causal directions between the two agendas.

Safety Valve or Pressure Cooker?

Findings:
Rather than simply being a “safety valve” that reduces bloggers’ anger over political or social problems, or a pressure cooker that increases these tensions, the results of this study suggested a more nuanced perspective in which the Chinese blogosphere can act as either, depending on the topic.

- **Safety value issues:**
  - Energy;
  - Politics;
  - Corruption and illegality;
  - Sino-Japan relations

- **Pressure cooker issues:**
  - Internet;
  - Religion;
  - Rural;
  - Arts;
  - Disasters.

Framing Effects based on N of Posts

Zhou & Moy (2007)

Yin (2010, in Chinese)
Summary of Empirical Evidence

What we know enough:
- Massive size of users
- Content censored or manipulated
- Active and influential opinion leaders

What we know little:
- Does the size matter?
- Is the censorship effective?
- Causality: Does social media bring social changes?
Discussion 1: State of the Art

Research on Chinese social media is

- a fast growing and increasingly popular field;
- prevailed by optimism by liberals, marketers, investors, and even the new generation of Chinese leadership;
- commonly examined with case studies, focusing on crisis events, controversial issues, or active groups;
- shortage of credible evidence for causal connections.
Discussion 2: Common and Old Problems

- The conceptual enthusiasm and empirical shortfall are not specific to research on Chinese social media;
- The same patterns occur in research on social media in general, probably at a slightly milder scale;
- These are also longstanding issues in research on traditional media worldwide;
- One common root of the above domains lies in the difficulty in getting necessary and quality data for the relevant research questions.
Price-Earning Ratio of Social Media

Sina
Tencent
Renren
Weibo
Facebook
Google
Twitter
Industry Average Price-Earning Ratio

Internet | Telecomm | Oil | Banking | Real Estate
---|---|---|---|---
China | U.S.
Discussion 3: Call for Smarter Research Design

- Need thoughtful and creative research designs, e.g.,
  - online experiments to ensure causal inference
  - individual-level analysis to ensure ecological validity
  - cross-national comparisons to enhance generalizability
Discussion 4: Opportunities and Challenges

Opportunities:
- Unique characteristics of the ecosystem behind GFW
- Active and sizable scholars in China for possible collaboration
- Increasing availability of open/affordable data for academic research

Challenges:
- How to engage in and contribute to mainstream research?
- How to find a dream collaborator with complementarity?
- How to evaluate data quality and interpret peculiar or counter-intuitive results?
Cited Works


THANK YOU & CONTACT US @

mailto: j.zhu@cityu.edu.hk
http://weblab.com.cityu.edu.hk
WB

Graph showing stock price movements from 2014/04/25 to 2014/07/25.
SINA Corp
NASDAQ: SINA - 25 Jul 3:59 pm ET

49.60 ↑ 0.15 (0.30%)

1 day 5 day 1 month 3 months 1 year 5 years max

Open 49.60
High 50.07
Low 49.08

Market cap 3.27B
P/E ratio (ttm) 139.12
Dividend yield -
Youku Tudou Inc (ADR)
NYSE: YOKU - 25 Jul 4:04 pm ET

19.69 📉 0.10 (0.51%)

1 day   5 day   1 month   3 months   1 year   5 years   max

Open  19.95
High  19.95
Low   19.25

Market cap  3.71B
P/E ratio (ttm) -
Dividend yield -

Google Finance - Yahoo Finance - MSN Money
Baidu Inc (ADR)
NASDAQ: BIDU - 25 Jul 4:00 pm ET

226.50 ↑22.23 (10.88%)

1 day 5 day 1 month 3 months 1 year 5 years max

Open 221.18 High 226.76
Low 217.00

Market cap 79.37B P/E ratio (ttm) 44.82
Dividend yield -

Google Finance - Yahoo Finance - MSN Money
Renren Inc
NYSE: RENN - 25 Jul 4:00 pm ET

3.30  \uparrow 0.02 (0.61\%) \\

<table>
<thead>
<tr>
<th>1 day</th>
<th>5 day</th>
<th>1 month</th>
<th>3 months</th>
<th>1 year</th>
<th>5 years</th>
<th>max</th>
</tr>
</thead>
</table>

Open 3.28
High 3.38
Low 3.20

Market cap 1.2B
P/E ratio (ttm) -
Dividend yield -

Google Finance - Yahoo Finance - MSN Money
Ctrip.com International, Ltd. (ADR)
NASDAQ: CTRP - 25 Jul 3:59 pm ET

67.00 ↑4.78 (7.68%)

1 day | 5 day | 1 month | 3 months | 1 year | 5 years | max

Open  62.56
High  67.23
Low   61.56

Market cap  9.07B
P/E ratio (ttm) 65.69
Dividend yield -
NetEase, Inc (ADR)
NASDAQ: NTES - 25 Jul 3:59 pm ET

83.63 ↓ 1.00 (1.18%)

<table>
<thead>
<tr>
<th></th>
<th>1 day</th>
<th>5 day</th>
<th>1 month</th>
<th>3 months</th>
<th>1 year</th>
<th>5 years</th>
<th>max</th>
</tr>
</thead>
</table>

- Open: 85.00
- High: 85.42
- Low: 83.45
- Market cap: 10.87B
- P/E ratio (ttm): 14.99
- Dividend yield: 0.41%

Google Finance - Yahoo Finance - MSN Money