Social Sensing and Analytics: 
Turning Big Data into Social Insights

Ee-Peng LIM 
Living Analytics Research Centre 
Singapore Management University
Social Media Everywhere
Twitter and Facebook

500 million users (Sep, 2013)
35.5 million users (China)
170 billion tweets
400 million tweets per day

1.15 billion users (Sep, 2013)
700 million daily active users
128 million daily active users (USA)
819 million monthly active mobile users
50 million Facebook pages
1.3 trillion Likes
What is Social Media?

- Users
- Network
- Content
Social Sensing and Analytics

Social Sensing

Data Base

Data Analytics

Applications
- CRM
- Recommendation
- Advertising
- Search
- Social policy
- Product marketing

Consumer insights
- Social insights
Example: Haze Event in Singapore (June 2013)

Daily Count of Tweets generated by ~150K Singapore users

www.channelnewsasia.com

www.todayonline.com
How much user attention was directed to haze?

Haze related tweets are ones containing #sghaze, #haze, #mustbehaze, and #blamethehaze.

Daily Haze-related **Tweet Count**

![Tweet Count Graph]

Daily Haze-related **User Count**

![User Count Graph]

14% new tweets tagged

More than 1/3 users
What happens to the emotion of users?

- **Anxiety**: “worried”, “worry”, “worries”, “fear”, “afraid”, “frightened”, “scared”, etc.
- **Low Arousal Negative**: “unhappy”, “miserable”, “sad”, “depress”, etc.
- **Gratitude**: “thank”, “thankful”, “grateful”, “blessed”, “lucky”, “fortunate”, “pleased”
- **Positive**: “happy”, “love”, “glad”, “pleased”, “relax”, “calm”, “relieve”, etc.
How closely people track NEA announcements?

NEA (National Environment Agency) Tweets on Jun 19:
• 1413 hrs 3-hour PSI is 170. Issued 2pm.
• 2102 hrs 3-hour PSI is 290. Issued 9pm.
• 2257 hrs 3-hour PSI is 321. Issued 10pm.

Retweet Count Per 10 Minutes for NEA Tweets
Challenges

Massiveness

Realtime

Data Stream

Noise
Palanteer Systems

Twitter Users from a Target community

Palanteer Systems

24/7 data crawling

Content
Tweets
Links

Search
Realtime Analytics
Event Detection
Consumer Intelligence

Computer Cluster
The Palanteer Systems
http://palanteer.sis.smu.edu.sg

Realtime Twitter Search

Realtime Social Analytics

Realtime Event Detection

Realtime Business Intelligence
Conclusion

- Social media = Platform for user interaction
- Social media = Source of social + consumer insights
- Massive, realtime, noisy, and stream social media data are difficult to tame

- LARC addresses the challenges by introducing social sensing and analytics research and system efforts.
The Future

Social Sensing Services

Data Analytics Services

Accelerated Generation of Consumer insights
Social insights

Applications
- CRM
- Recommendation
- Advertising
- Search
- Social policy
- Product marketing

Accelerated Deployment of Insights
This work is contributed by the LARC’s online community research team

LARC is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office