Building a Social Media Tool
for Sociological Studies – BlogTrackers

Huan Liu

Joint Work with Shamanth Kumar and Ali Abassi

http://dmml.asu.edu/

SciTechBlog

May 21, 2010

Does Facebook know who you'll date next?

Posted: 01:29 PM ET

Facebook co-founder Mark Zuckerberg genuinely cares about your relationship status. Well, maybe not genuinely, but he may look into it if he's bored.

One of Zuckerberg's favorite pastimes is determining Facebook users' relationship patterns, according to David Kilpatrick's upcoming book, "**The Facebook Effect**: The Inside Story of the Company That Is Connecting the World."

Roles of Blogs in Social Media

- Citizen journalism
- Greater freedom and protection to express opinions
- Blogs are a more scalable alternative to traditional methods of opinion mining, like surveys
- Blog sites are potentially rich sources of Information because of their diverse content
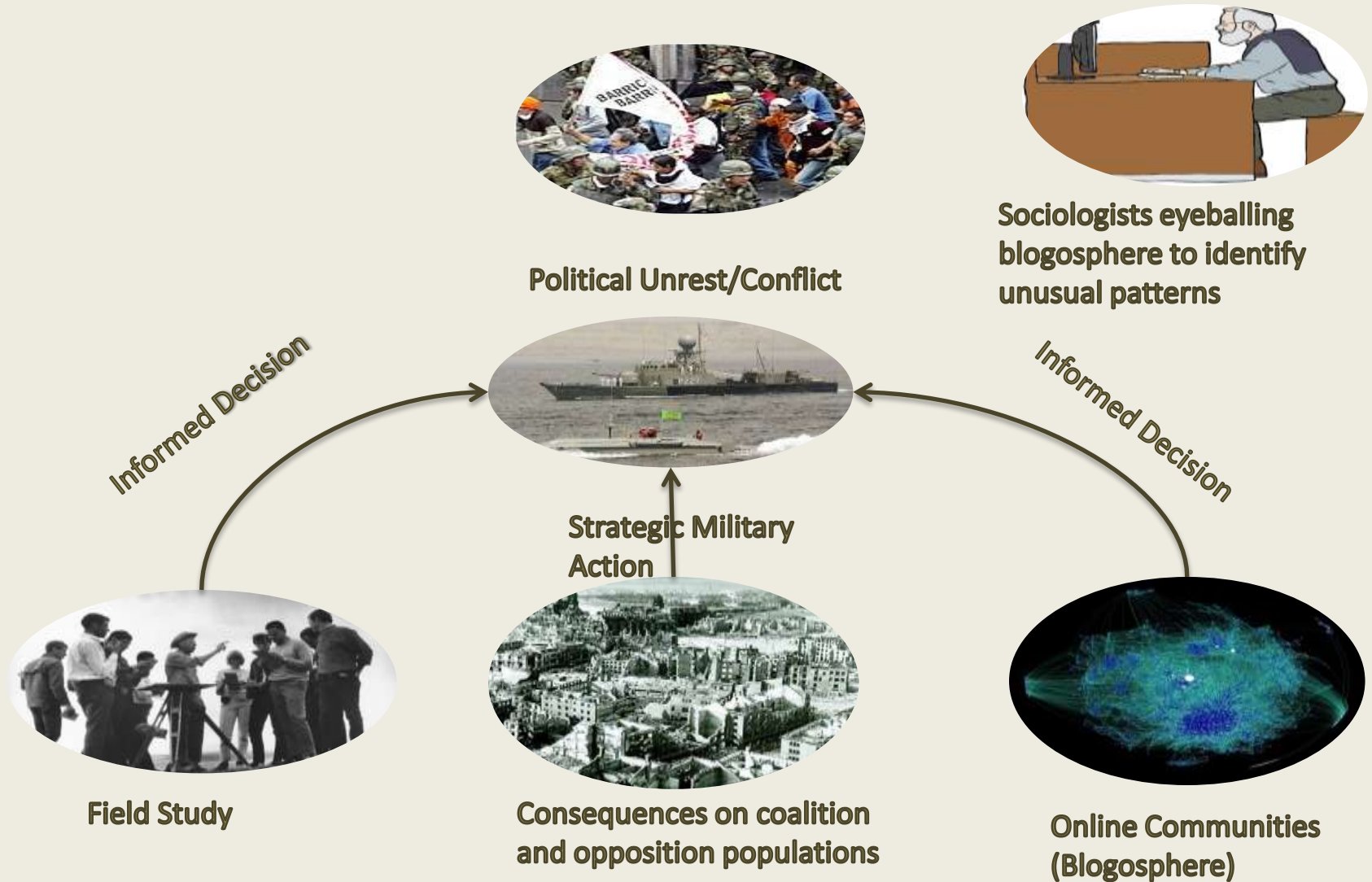
Key Features of BlogTrackers

- Influential bloggers: bloggers with the capability to influence a significant number of other individuals. Influential bloggers represent notable members of the community. The topics these members discuss are likely to be more important than a random blogger
- Traffic Pattern Graphs: to detect events through traffic anomalies
- Future reference tracking of keywords through e-mail
- Domain specific phrase detection through keywords

A Challenge

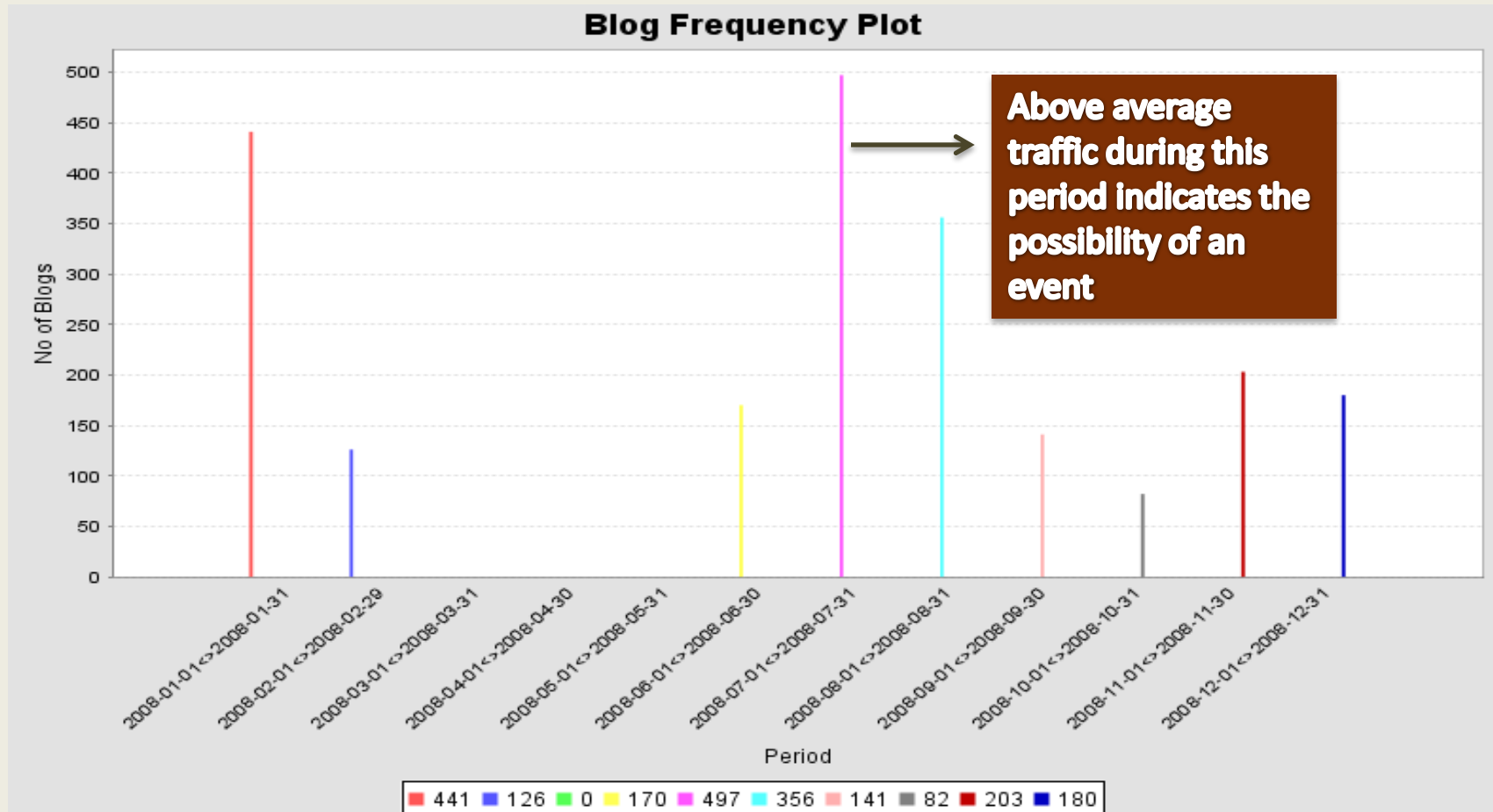Political Unrest/Conflict

Sociologists eyeballing blogosphere to identify unusual patterns

Informed Decision

Informed Decision

Strategic Military Action

Field Study

Consequences on coalition and opposition populations

Online Communities (Blogosphere)

A Typical Blog Analysis Workflow

Select data source: blog sites of particular interests

↓

Identify events of interest

↓

Understand events from the blog posts and observe post event reactions

↓

Identify influential bloggers or opinion leaders in the community

↓

Observe future references to the events

Traffic Pattern Graphs

Blog Frequency Plot

Above average traffic during this period indicates the possibility of an event

Watch-Lists and Alerts

Online e-mail notification.

Keyword [　　　　　　　　　|　　　　　] [Add]

☑ Enable E-mail notification [　　　　　　　　　] [Add] [Delete]

Current E-mail IDs   shamanth00@gmail.com,

Current Watch List
ces
ipod
touch
breakingnews
macworld

[Delete]

Tags	
iphone	
ces	
macworld	
ipod	
store	
touch	
app	

[Insert]

Blogger	BlogPost
Darren Murph	Acoustic Research Infinite...
Paul Miller	Ulysse Nardin's Chairman ...

Offline notification of posts containing the watch list keywords.

System generated popular keywords

Ongoing Work and Acknowledgements

- Web version of the tool for easier access and expertise shsaring
- API release to enable expert comments/critiques
- Open Source migration
- Additional forms of analysis, e.g., network & sentiment analysis
- Expanding our blog sources

This work is supported, in part, by the Office of Naval Research (ONR). Images were sourced from Google Image Search. We also thank Dr. Merlyna Lim (School of Justice and Social Enquiry) for her input.

BlogTrackers in Use

Online social behaviors and prediction of their implications for the physical world

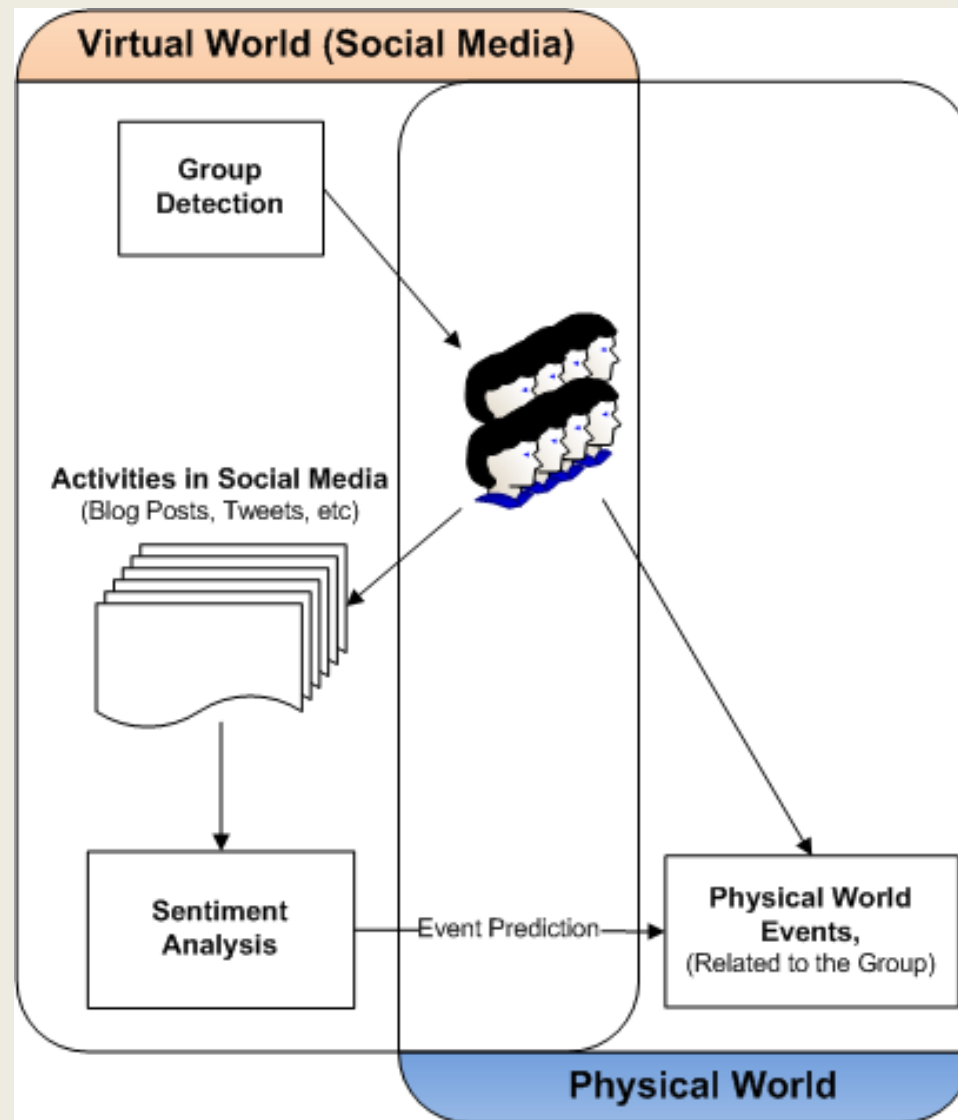- A Comparative Study (HSCB Project with UH)

Project Overview

Goals

➢ Studying the attitudes of individuals and groups based upon the online content (a virtual world)

➢ Understanding and anticipating actions in the real world
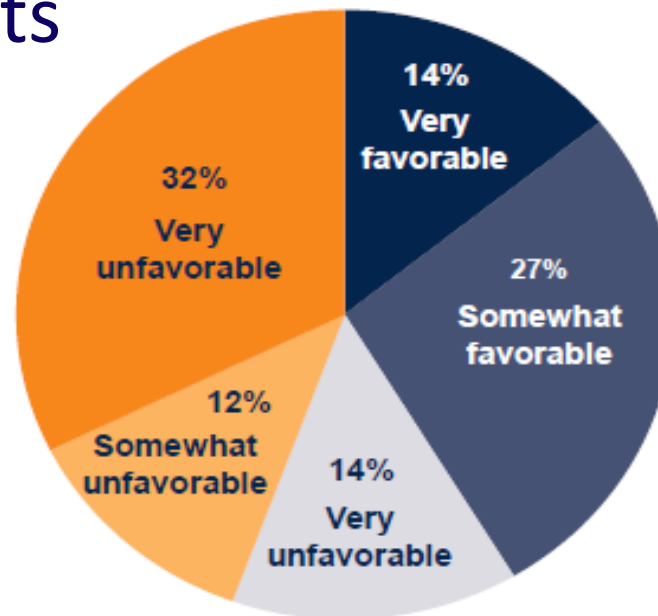
Tasks

1. Social Science Research for Comparative Study of Attitudes
2. Research about Group and Behaviors on Social Media
3. Theory-Guided, Iterated Model of Group Behavior

Virtual World (Social Media)

Group Detection

Activities in Social Media
(Blog Posts, Tweets, etc)

Sentiment Analysis

Event Prediction

Physical World Events,
(Related to the Group)

Physical World

Sample- Kaiser's Health Tracking poll

- Is a monthly polling report, talks about the trend of opinions regarding the health issues.
- And provides variety of health care opinion polling results

CHART 1

Views of New Health Reform Law

As you may know, a new health reform bill was signed into law earlier this year. Given what you know about the new health reform law, do you have a generally favorable or generally unfavorable opinion of it?

- 14% Very favorable
- 27% Somewhat favorable
- 14% Very unfavorable
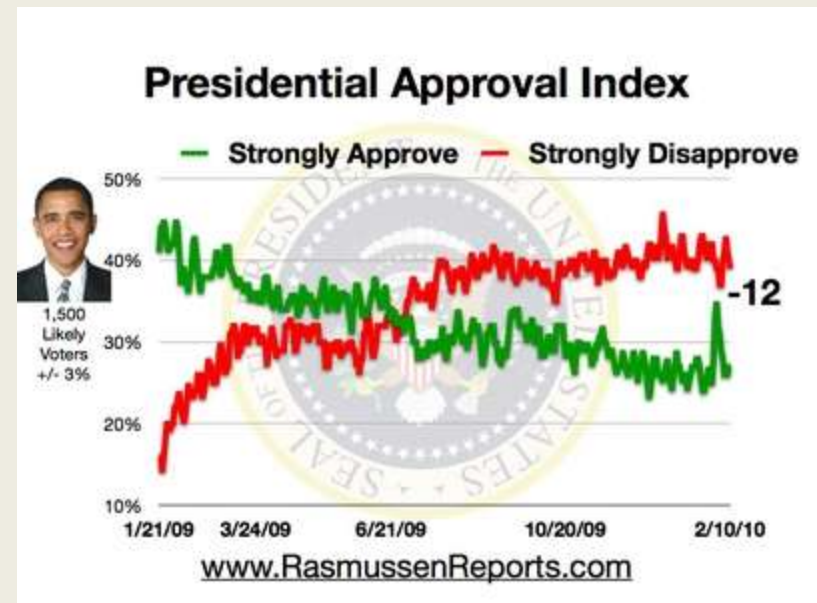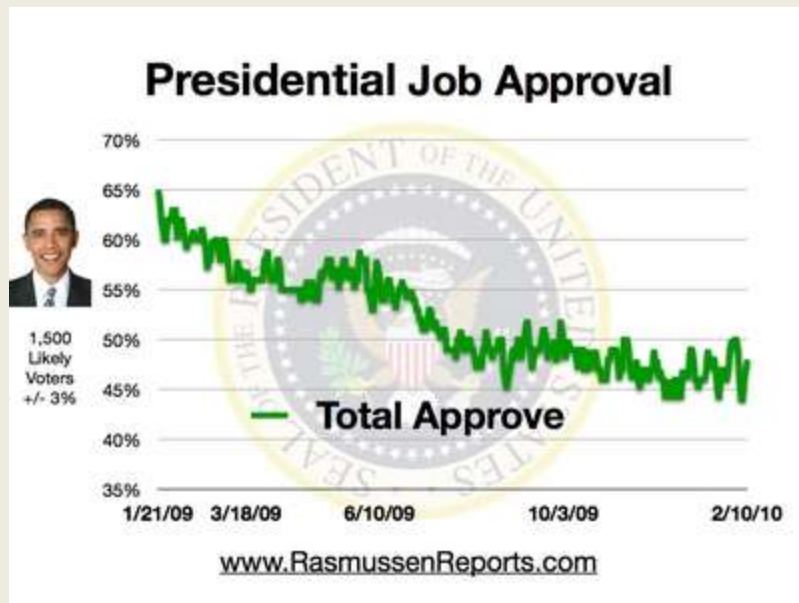- 12% Somewhat unfavorable
- 32% Very unfavorable

Source: Kaiser Family Foundation *Health Tracking Poll* (conducted May 11-16, 2010)

Sample- Presidential approval index

- Rasmussen Reports
  - Daily Presidential Tracking Poll
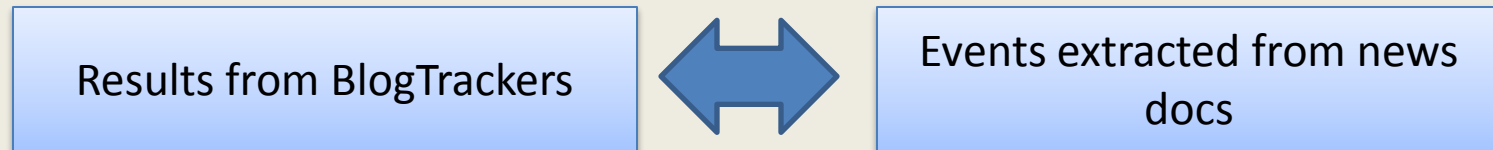 - Trend

Working with BlogTrackers Data

- Event Extraction

  – Major and minor events

- Finding influential bloggers

- Finding initiators' documents

- Model document and idea propagation in weblogs (or how an idea or request spreads)

Data Analysis

- Finding the relationships between
  - Variables representing patterns of blogging activity in a country at particular points in time
  - Variables representing actual political events during those same periods, as well as before and after

- Evaluation

| Results from BlogTrackers | ⟷ | Events extracted from news docs |

- Validation from the filed

Challenges: Measures and Comparisons

To quote Sun-Ki's message to us

- When data is available on both the content and network connections of units of analysis, how do the two types of information interact in developing metrics for prediction?

- What sorts of methods are most suitable for determining the reliability and validity of used measures?

- To what extent, and under what conditions, can analysis of offline and online networks complement one another? How well can empirical findings for offline networks inform analysis of online networks?

Conclusions and Acknowledgments

- Many questions remain - interdisciplinary efforts
- Need theories to guide computation-based investigation
- Integrate social theories and social computing
  - To push the boundary of social sciences
  - To investigate hypotheses to understand human behavior in social media

Acknowledgments

This is a joint project with University Hawaii sponsored by ONR HSCB (Sun-Ki Chai)

Additional Information

- Social Computing, Behavioral Modeling, and Prediction
  - sbp.asu.edu
  - A platform to exchange and collaborate
- SBP11 will be held in Maryland
- A book on Modeling and Data Mining in Blogoshpere