The Web: Humanity’s Sensor Network

Jure Leskovec (@jure)
Including joint work with A. Anderson, D. Huttenlocher, C. Danescu-Niculesci-Mizil, D. Jurafsky, J. Kleinberg, S. Myers, C. Potts, D. Shahaf
Web – Diverse Applications
Web – Rich Interactions

Rich & dynamic interactions between users and the content.
Web – Large Scale

Large scale.
Billions of users.
Web – Traces of Activity

Massive digital traces of human activity.
The Web is our “laboratory” for understanding the pulse of humanity.
Invisible Becomes Visible

Science advances when invisible becomes visible
Recognizing Behavior

Can we recognize fundamental patterns of human behavior from raw digital traces?
Digital Traces → Networks

Model as an interaction network
To Make Progress We Need

1) New branch of mathematics:
   CS theory and graph theory

2) New algorithms:
   Data mining and machine learning

3) New computing platforms:
   Large-scale distribute data processing
People and Communities

How do people become members of communities?
Communities and Language

- Language builds collective identity
- Linguistic change captures the relation between a member and the community
Mysteries of language change

- Innovations come from adolescent (age 16-19) women
  Why? Biological critical period?
  Teenage rebellion?

- Stability (freezing) in language of older speakers
  Why? Biology? Are adults really stable?
Obtaining Linguistic Record

Follow people with microphones for 50 years. Or, get the complete linguistic record (10 years) of an online community:

I had a cluster of microcalcifications in my rt breast that they have been following. Went in to have a wire loc because the radiologist said that a stereotatic wasn't possible. We'll the radiologist was unable to wire loc (didn't even try the day of surgery), b/c they are faint and difficult to obtain. So I've been having 6 month f/u mammo and MRIs.

Clear copper colored brew, medium cream colored head. Floral hop nose, caramel malt. Caramel malt front dominated by a nice floral hop background. Grapefruit tones. Very tasty hops run the show with this brew. Thin to medium mouth. Not a bad choice if you're looking for a nice hop treat.
“Aroma” vs. “Smell”

Convention usage over years for 'Aroma' and 'Smell' conventions.
Young Adopt Innovations

Users who joined in 2003

Users who joined in 2005
Community/Individual Change

Individual:

“life stage”

Community:
Increased Identification

Increased identification with the community
Fruit words (peach, pineapple, berry, ...) are getting ever more popular
Member-Community Change

Community language

Member

Distance from the community

Time

Member Life-stage
Member-Community Change

- Member
- Community language
- Time
- Member Life-stage
- Distance from the community

Graph showing the change in community language over time with respect to member life-stage.
User-Community Change

Before leaving, users stop adapting
Elastic Life-cycle

Given the first 20 posts, we can predict whether the user is about to leave the community
How to steer human behavior?
People work amazingly hard to earn badges
“Give me enough medals and I’ll win you any war.”
— Napoleon
Scientific Communities

Natural analogies to how evaluation works in scientific communities:

- Acceptance of papers to conferences and journals.
- Funding of grant proposals.
- Who gets hired, who receives awards, ...

A lot of decentralization, but to some extent a planned economy:

- Program/hiring/award committees, ...
- Strategic priorities of funding agencies.
- Communities articulating their own strategic priorities.

Badges play multiple roles:

- Recognize a wide range of types of activities
- Serve both as credentials and create incentives
Badges & Behavior Change

How do badges translate into effects on user behavior?

Need lots of data to tease out the effects and build a mathematical model

Badges on Stack Overflow Q&A site:
2M people
5M questions
10M votes

How to format a JSON date?

I'm taking my first crack at Ajax with jQuery. I'm getting my data onto my page, but I'm trouble with the JSON data that is returned for Date data types. Basically, I'm getting looks like this:

```json
/Date(1224043200000)/
```

From someone totally new to JSON - How do I format this to a short date format? She handled somewhere in the jQuery code? I've tried the `jQuery.ui.datepicker.p3.datepicker.formatDate()` without any success.

FYI: Here's the solution I came up with using a combination of the answers here:

This solution got my object from the callback method and displayed the dates of date format library.
2 parts to the model:

User gains value from obtaining a badge
But it “hurts” user to change behavior
Badges are “Magnets”
User **steers** in badge direction as she approaches it; then **resets** after receiving it.
Vote on 600 questions
Question: How should you “place” badges to achieve desired effects?

Our method can find **optimal** badges:

![Graph showing change in behavior vs. badge position](image)
Badges in Massive Online Courses

Machine Learning

Statement of Accomplishment

August 20, 2012

Dear Marta

Congratulations! You have successfully completed the online Machine Learning course (ml-class.org), offered April through July, 2012. To successfully complete the course, students were required to watch lectures, review questions, and complete programming assignments. Your score was as follows:

A class of 110,000 students
Badges for motivating forum contributions

**The Reader**
To earn the next badge (Silver), you must read 30 threads from your classmates.

**The Supporter**
To earn the next badge (Silver), you must vote on 15 posts that you find interesting or useful.

**The Contributor**
To earn the next badge (Bronze), you must post 3 replies that your classmates find interesting.

**The Conversation Starter**
To earn the next badge (Bronze), you must start 3 threads that your classmates find interesting.

**Top Posts**
To earn the next badge (Bronze), you must write a post that gets 5 upvotes from your classmates.
Effects of Badges

Badges lead to increased engagement:
Reading and voting but not posting

3 runs of a course:
- No
- badges
- Badges

(a) Number of votes
(b) Number of reads
Observations: Data analysis

Algorithms: Applications

Models: Predictions

Actively influencing the system

Massive data: Observe “invisible” patterns
THANKS!
http://snap.stanford.edu
@jure
References
