Analysis of Online Networks

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

Diverse on-line computing applications.

Web – Rich Interactions

Rich & dynamic interactions between users and the content.

Web – Large Scale

Large scale.
Billions of users.

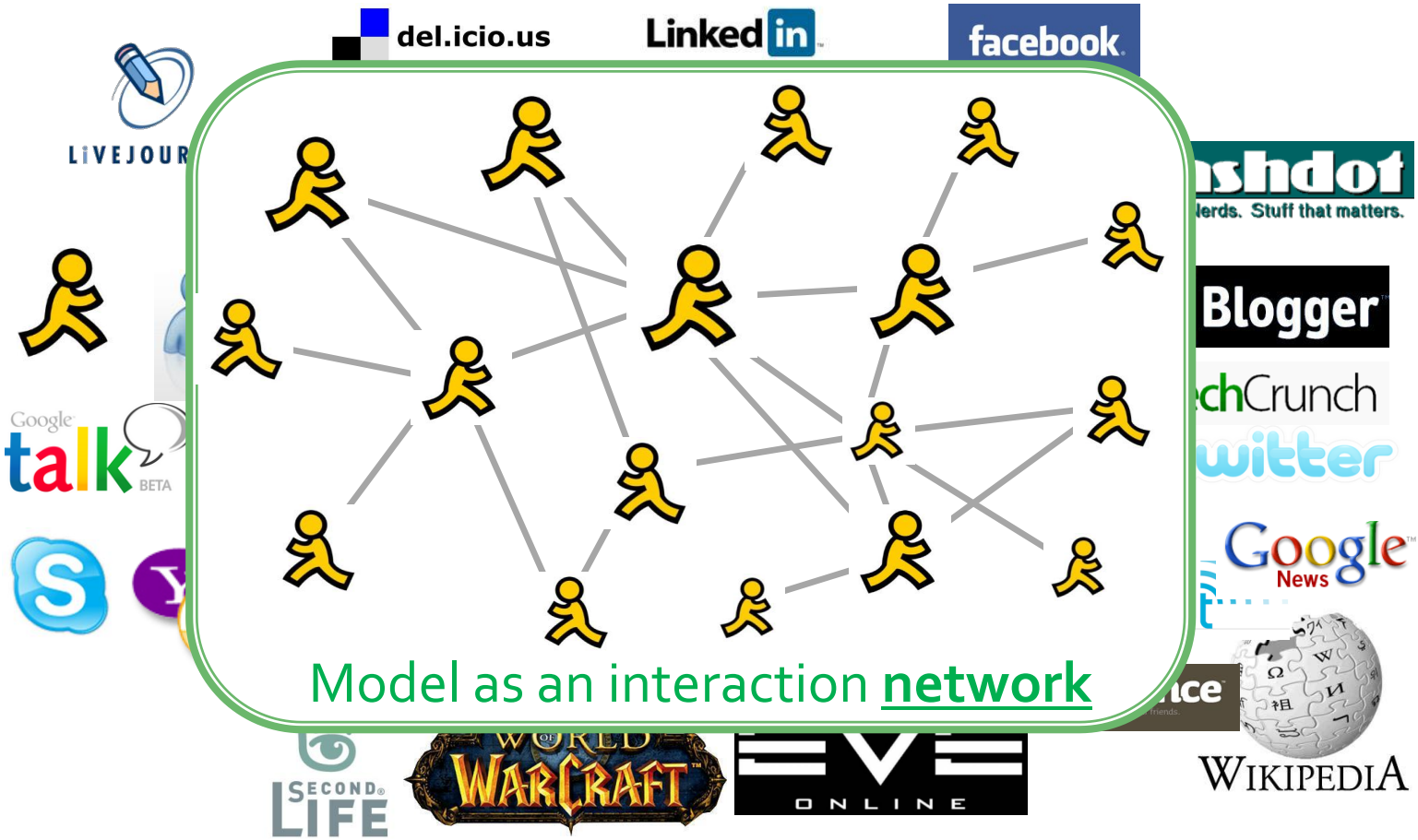
Web – Massive Traces

Massive digital traces of human activity.

Web – Lab for Humanity

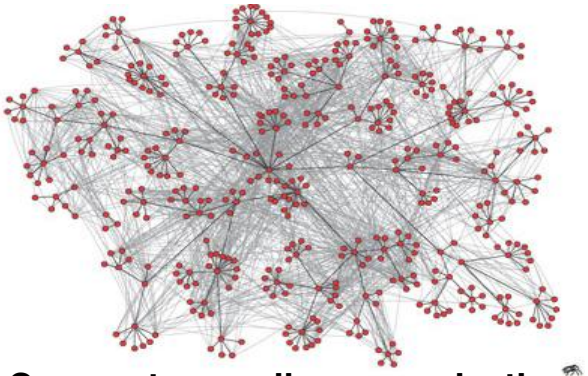
The Web is our "laboratory" for understanding the pulse of humanity.

Digital Traces → Networks



Transformation of Computing

Online friendships
[Ugander-Karrer-Backstrom-Marlow, '11]



Corporate e-mail communication
[Adamic-Adar, '05]

Web is a sensor into humanity!

Profound transformation in:

- How knowledge is produced and shared
- How people interact and communicate
- The scope of CS as a discipline

Networks: Size Matters

Network data: Orders of magnitude

- 436-node network of email exchange at a corporate research lab [Adamic-Adar, '03]
- 43,553-node network of email exchange at an university [Kossinets-Watts, '06]
- 4.4-million-node network of friendships on a blogging community [Liben-Nowell et al., '05]
- 240-million-node network of communication on MSN Messenger [Leskovec-Horvitz, '08]
- 800-million-node network of declared Facebook friendship [Ugander et al., '11]

My group's research

New methods

- Graph and Data mining
- Large scale machine learning

New tools

- Dynamics and flow of online information
- Taming information overload

New science

- Recognize fundamental human behaviors
- Influence human behavior

Patterns of Human Behavior

Science advances when invisible becomes visible

Social interaction is leaving digital traces on-line

Can we recognize fundamental patterns of human behavior from raw digital traces?


Image credit: Google Image Search

Patterns of Human Mobility

[KDD '11]

Mobility & Social Networks

What's the relation between person's mobility and her network?

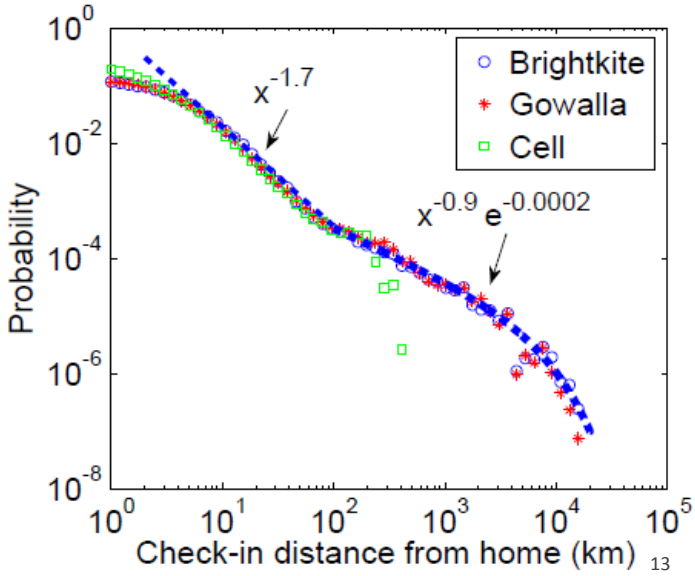
Location-based online social networks

Brightkite, Gowalla: 10m check-ins

Cell phones

Portugal: 500M calls

In terms of mobility the datasets are indistinguishable!



[KDD '11]

Modeling Human Mobility

Towards a model

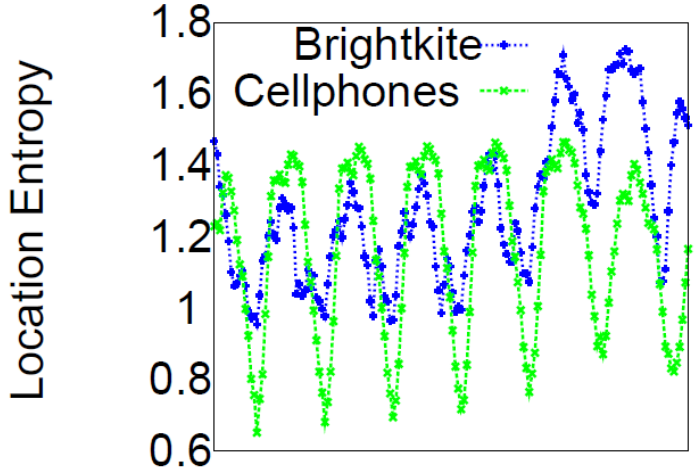
Observations:

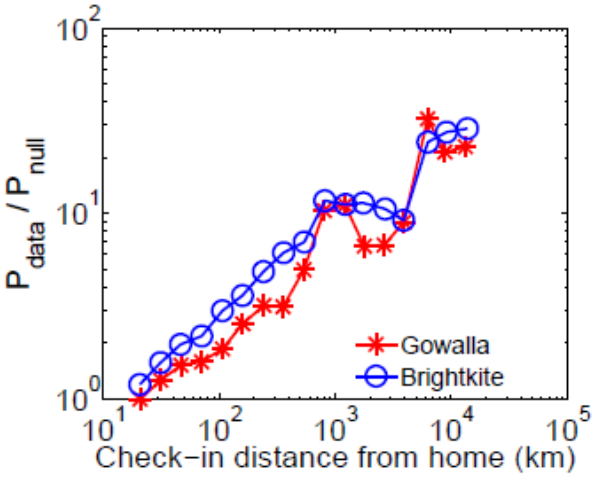
High spatial and temporal periodicity:

- 30-53% of locations user visits multiple times
- Low location entropy at night/morning
- Higher entropy during the day

Social network effects:

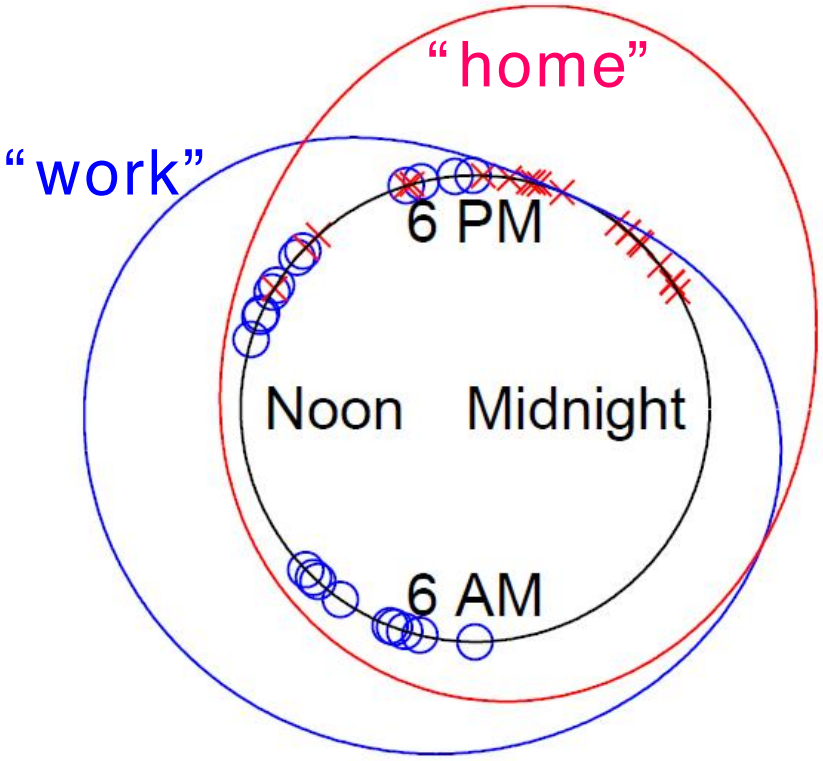
- Especially over the weekend





[KDD '11]

2 Components to the Model

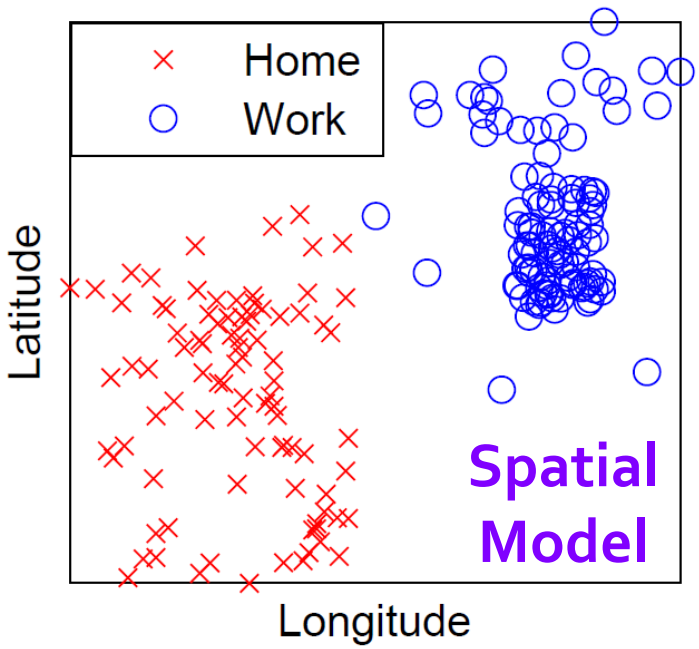


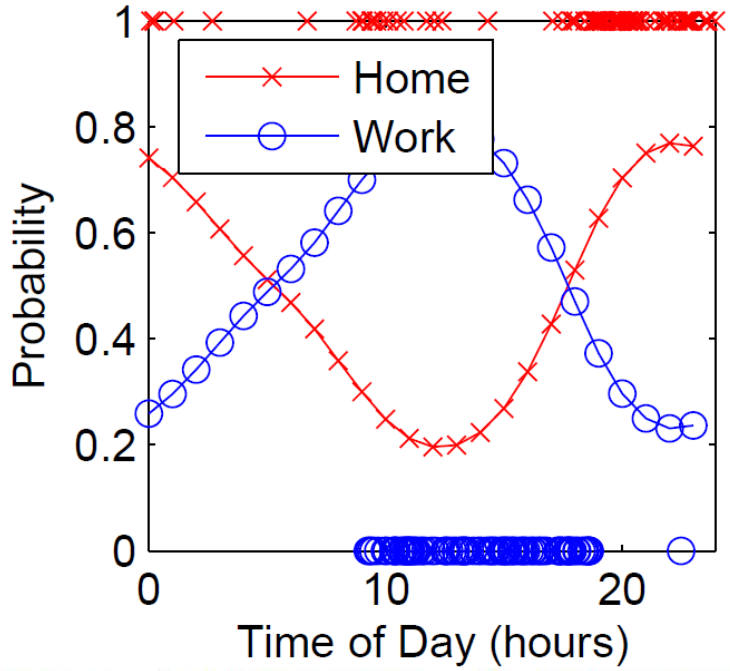
Spatial model:
"Home" vs. "Work"

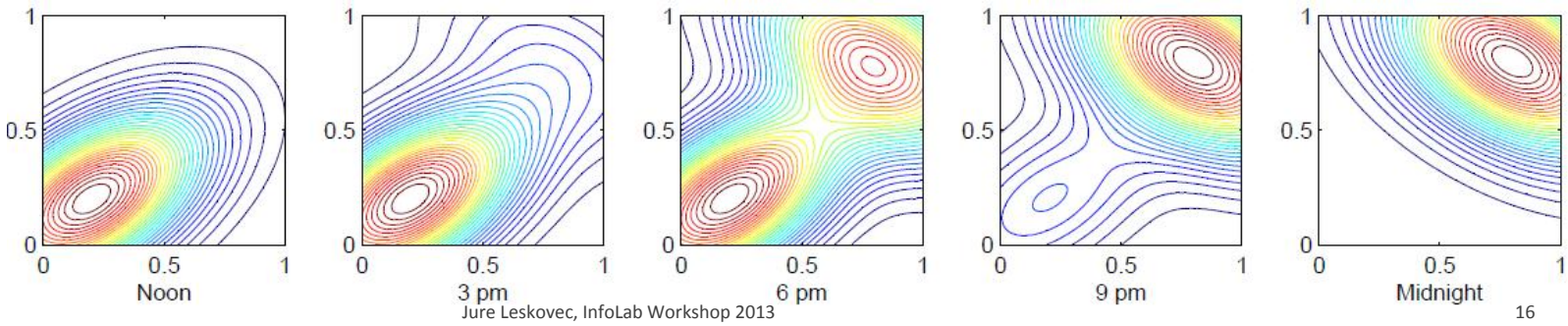
Temporal model:
Commuting around home/work

[KDD '11]

Example: A Gowalla User







[KDD '11]

Weekend Mobility

Social network plays particularly important role on weekends

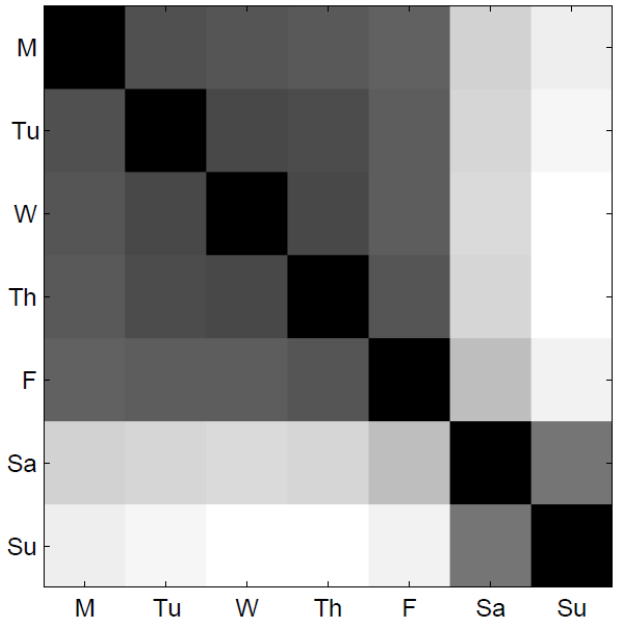
Include social network into the model

Prob. that user visits location X depends on:

1) Distance(X, F)

2) Time since a friend was at location F

F = Friend's last known location

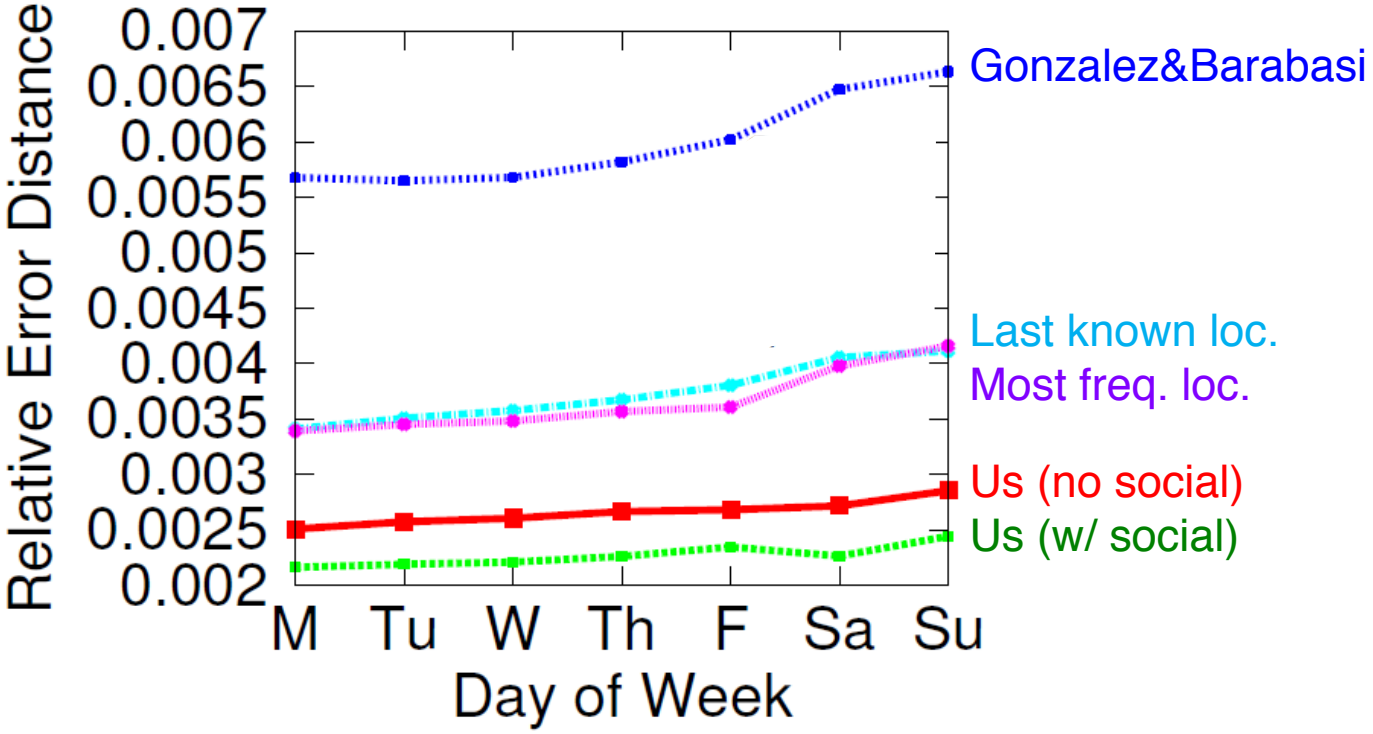


Mobility similarity

[KDD '11]

Mobility: Results

Cellphones: Whenever user receives or makes a call predict her location



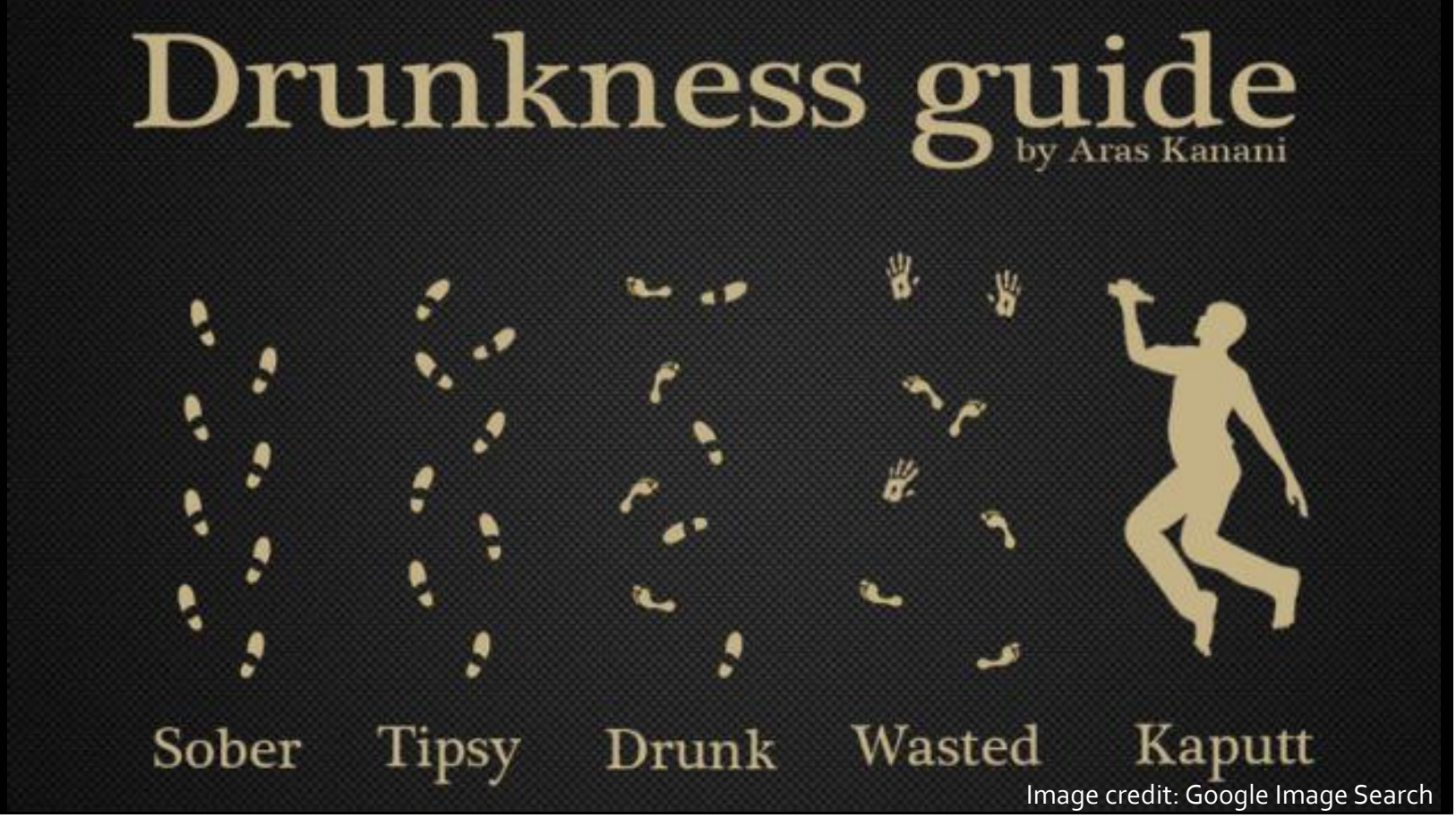


Image credit: Google Image Search

How people become members of communities?

Online Communities

People discuss beer:

Ba

Tballz420

4/5 rDev +1.8%

look: 3 | smell: 4 | taste: 4 | feel: 3.5 | overall: 4.5

Clear copper colored brew, medium cream colored head. Floral hop nose, caramel malt. Caramel malt front dominated by a nice floral hop backround. 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.

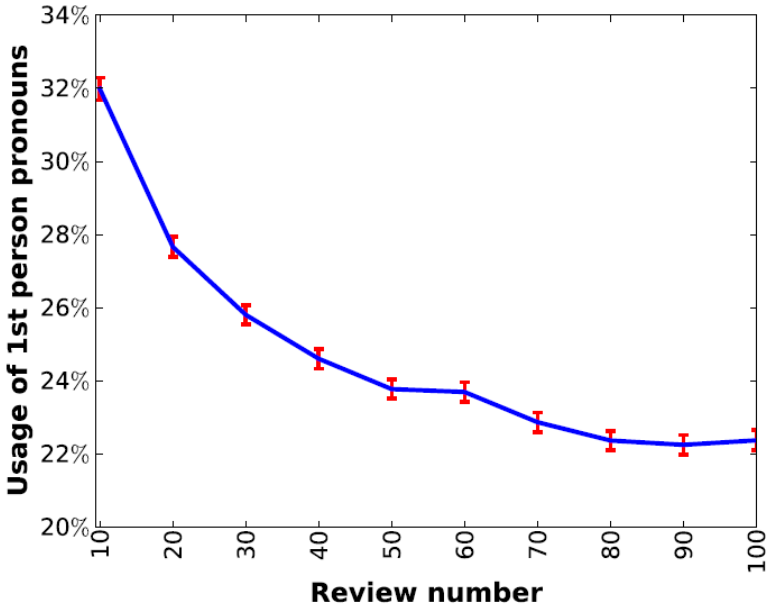
BeerAdvocate.com

ratebeer

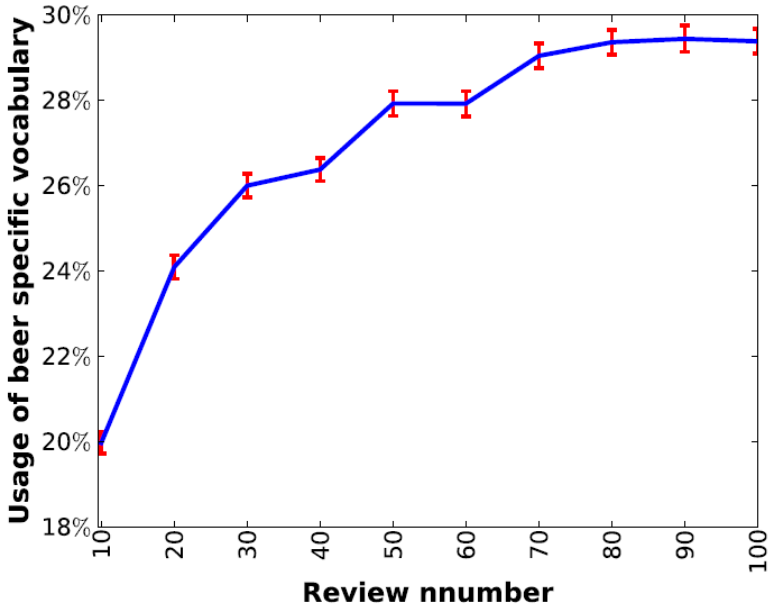
Q1) How does user's language evolve as they get more into the community?

Q2) How does it change just before they leave the community?

User Language Change



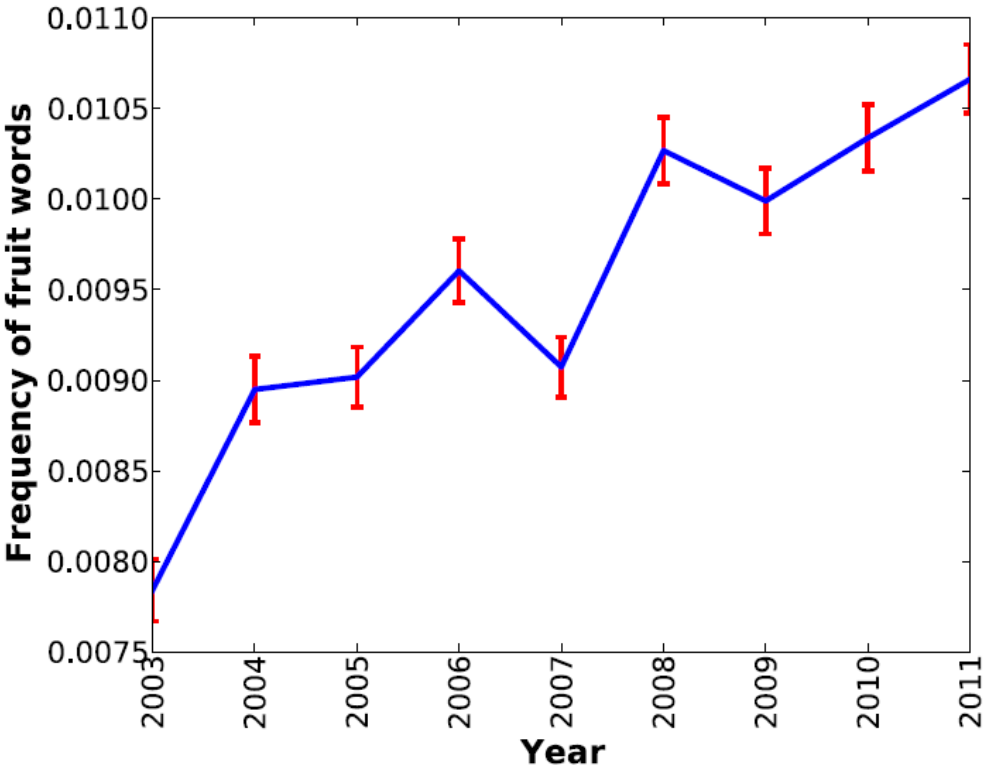
(a) First person sing. pronouns



(b) Beer specific vocabulary

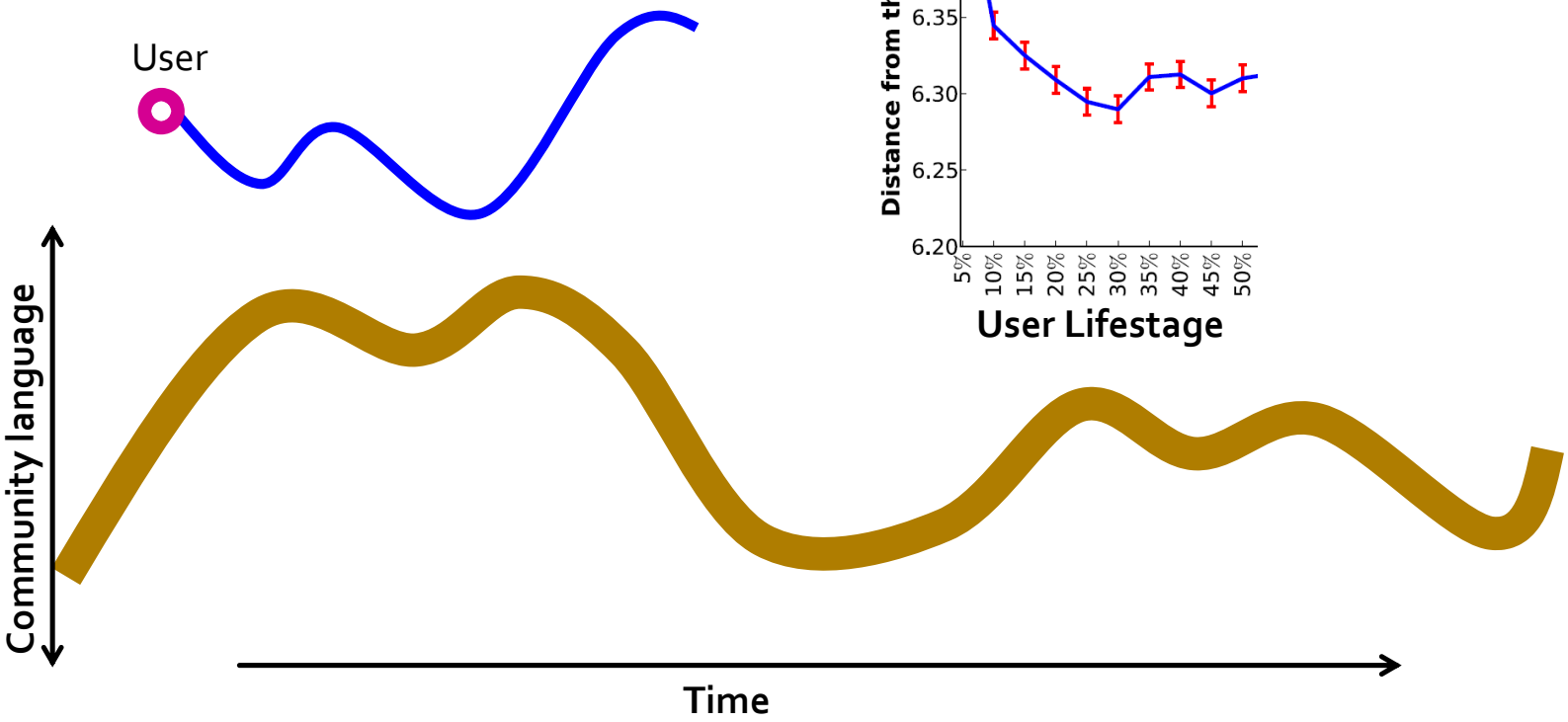
User's use less first person pronouns but more beer specific language

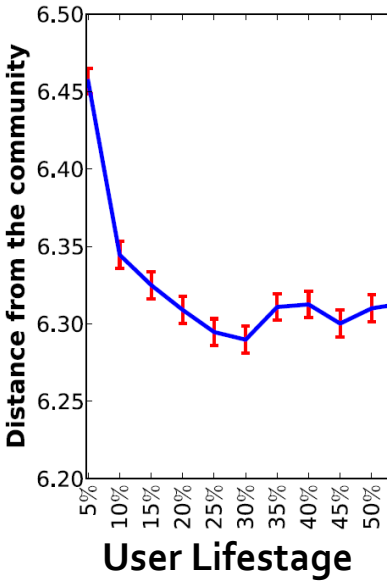
Community Language Change



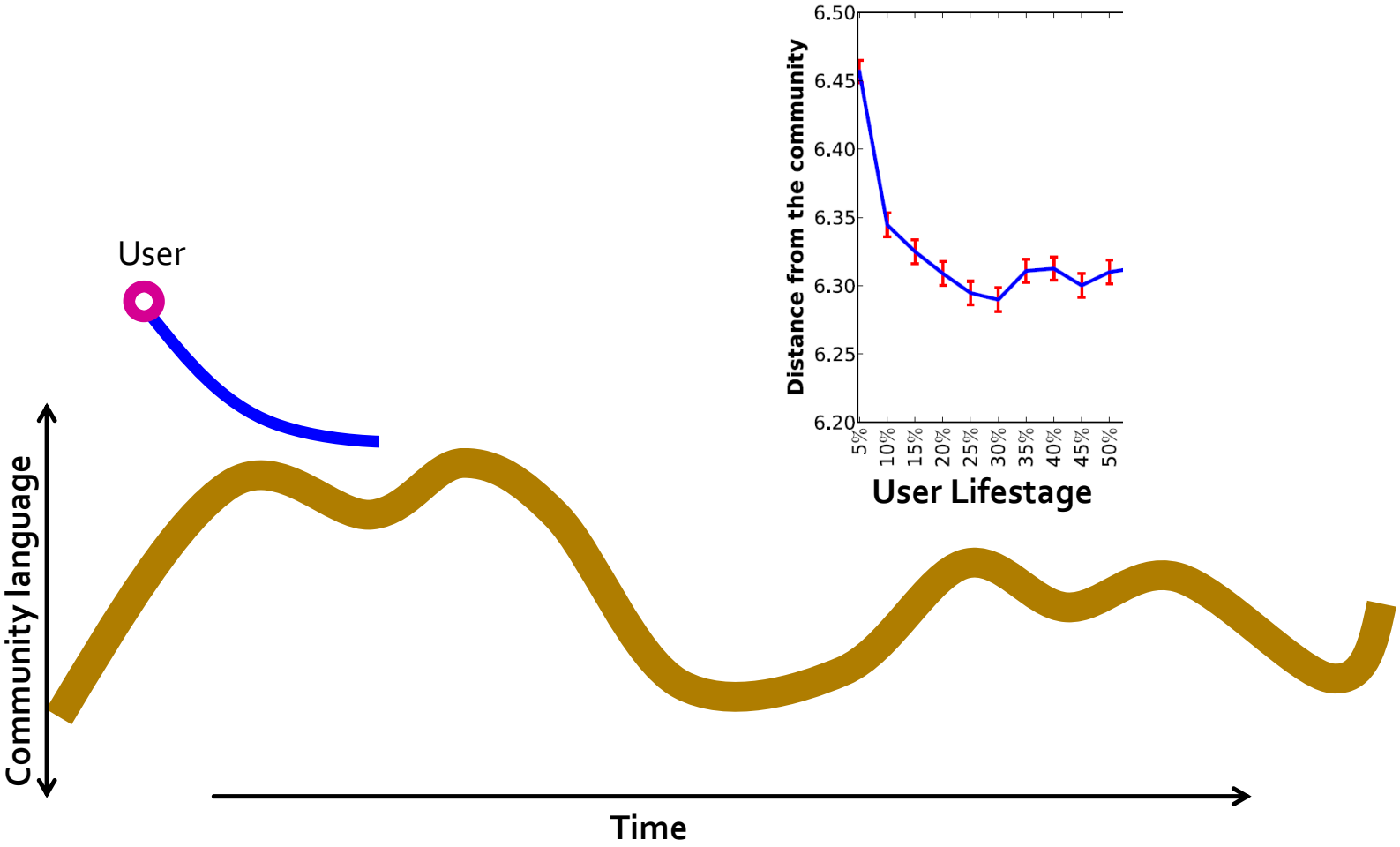
Fruit words (peach, pineapple, berry, ..) are getting ever more popular

User-Community Change

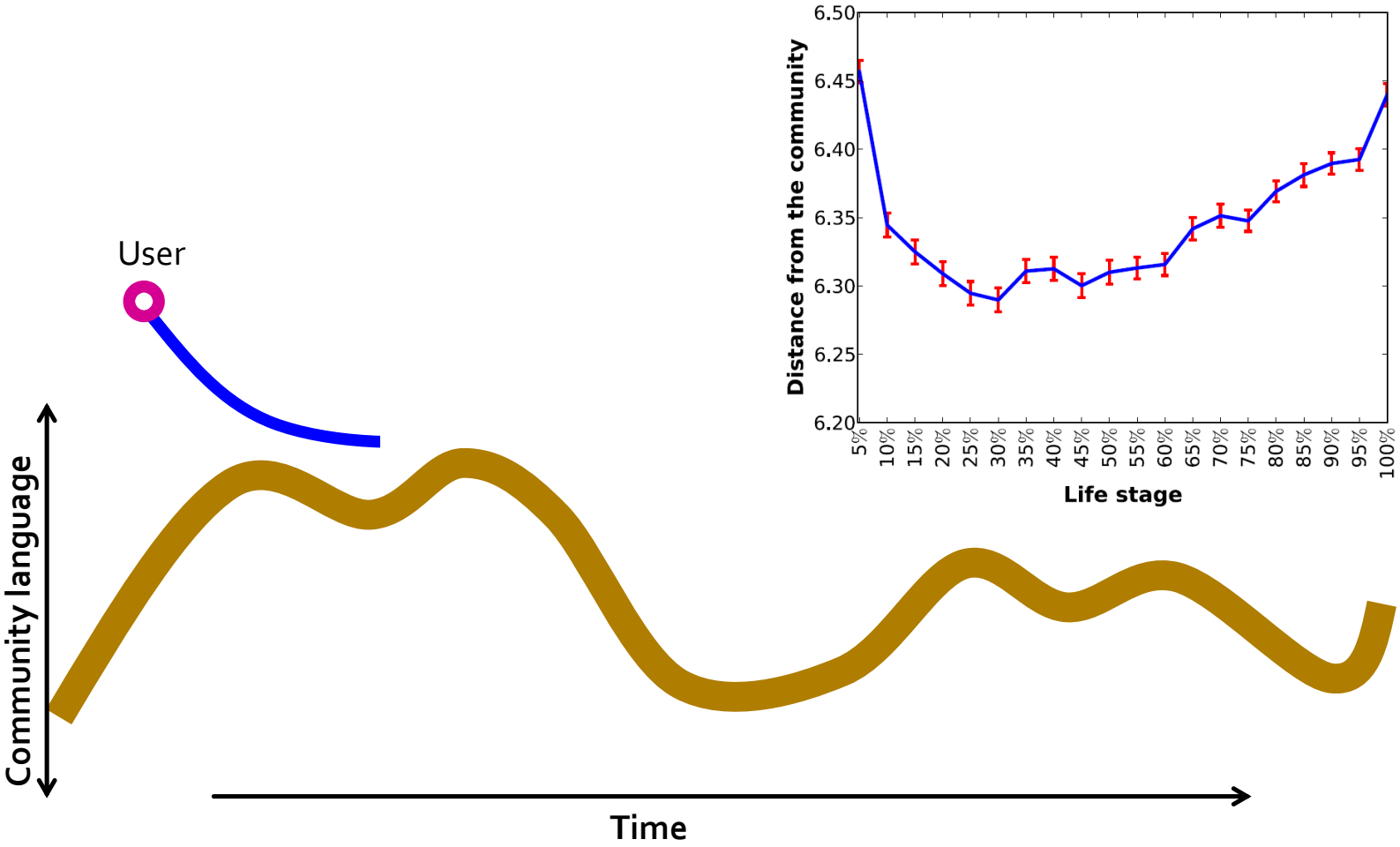




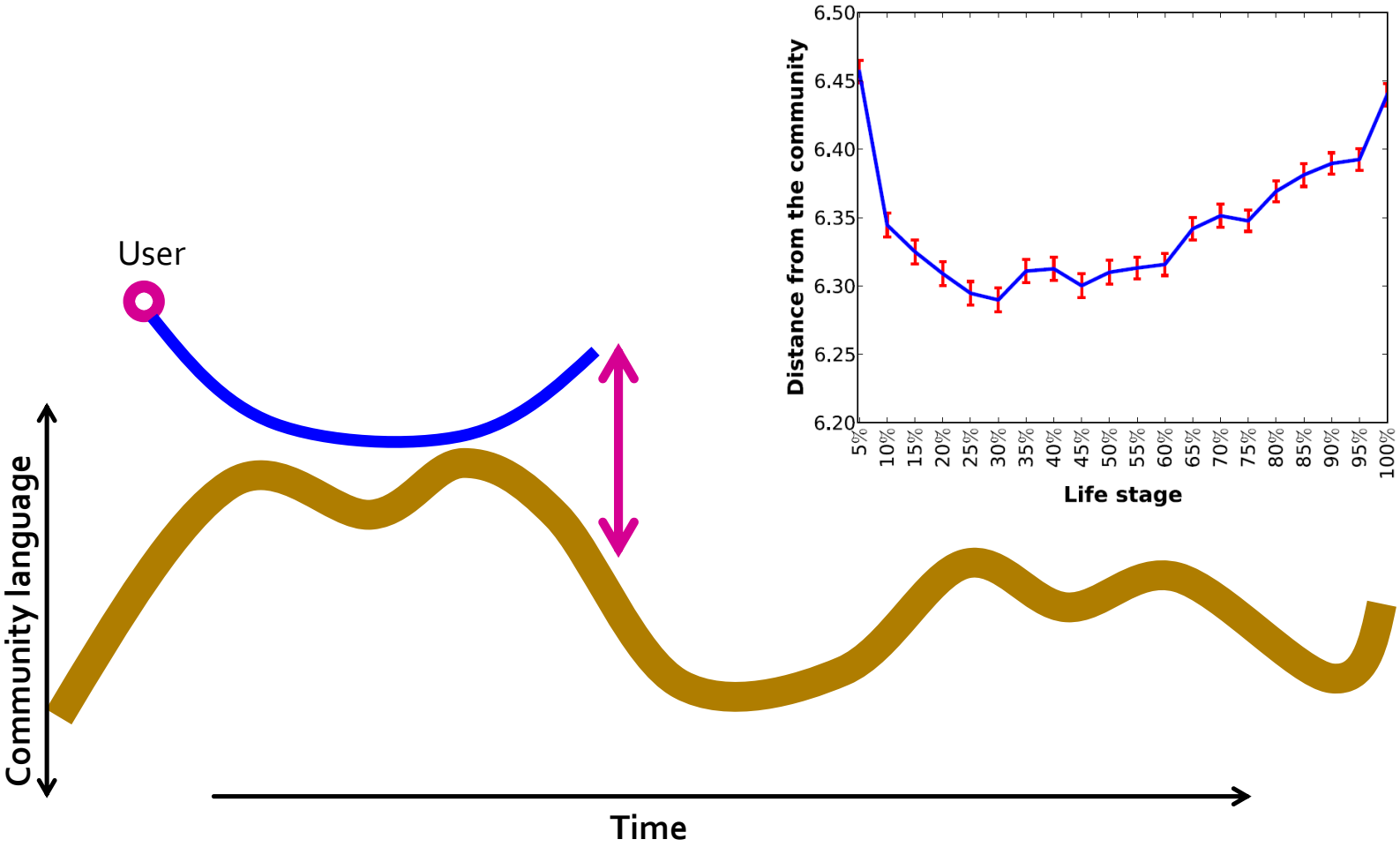
User-Community Change



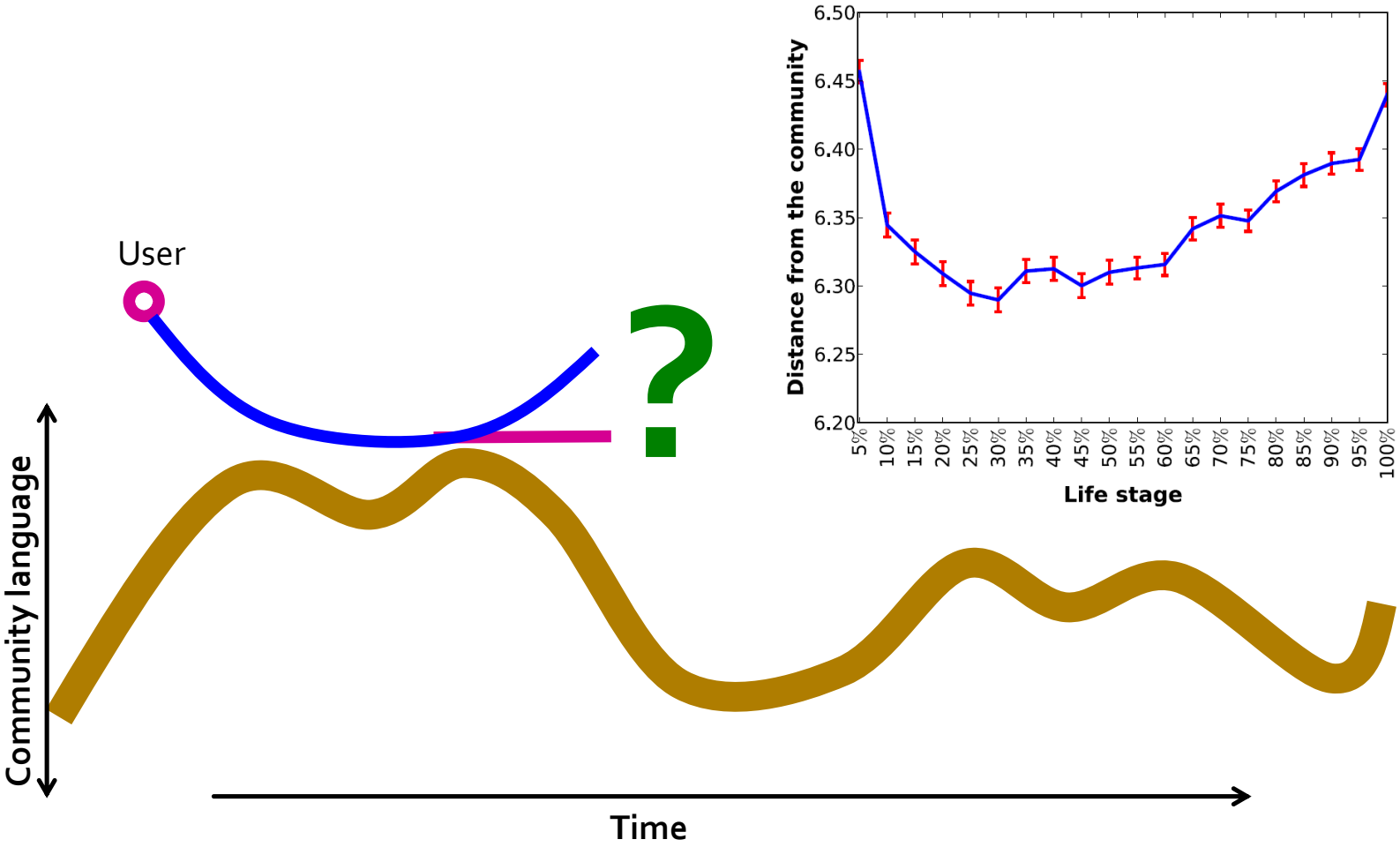
User-Community Change



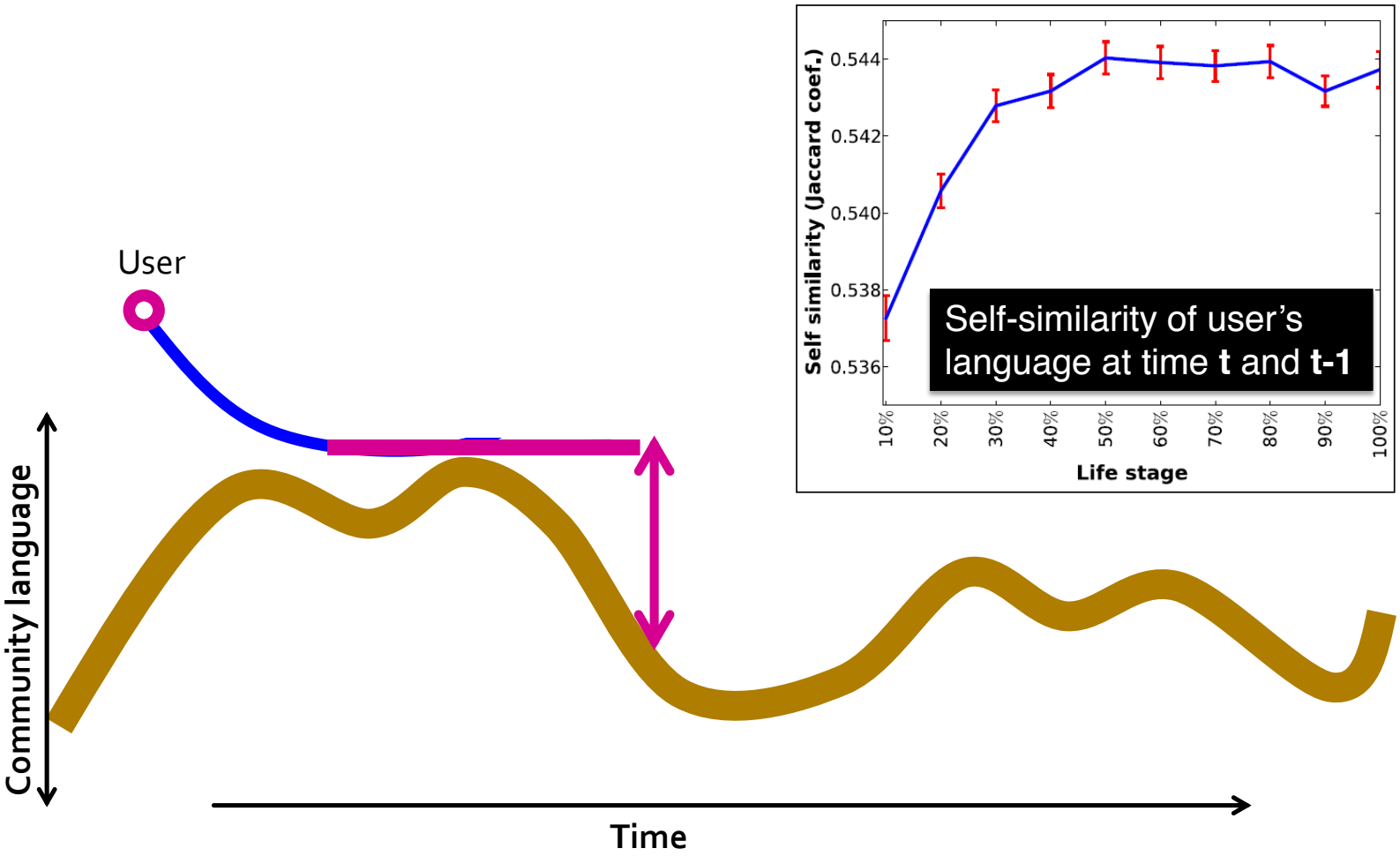
User-Community Change



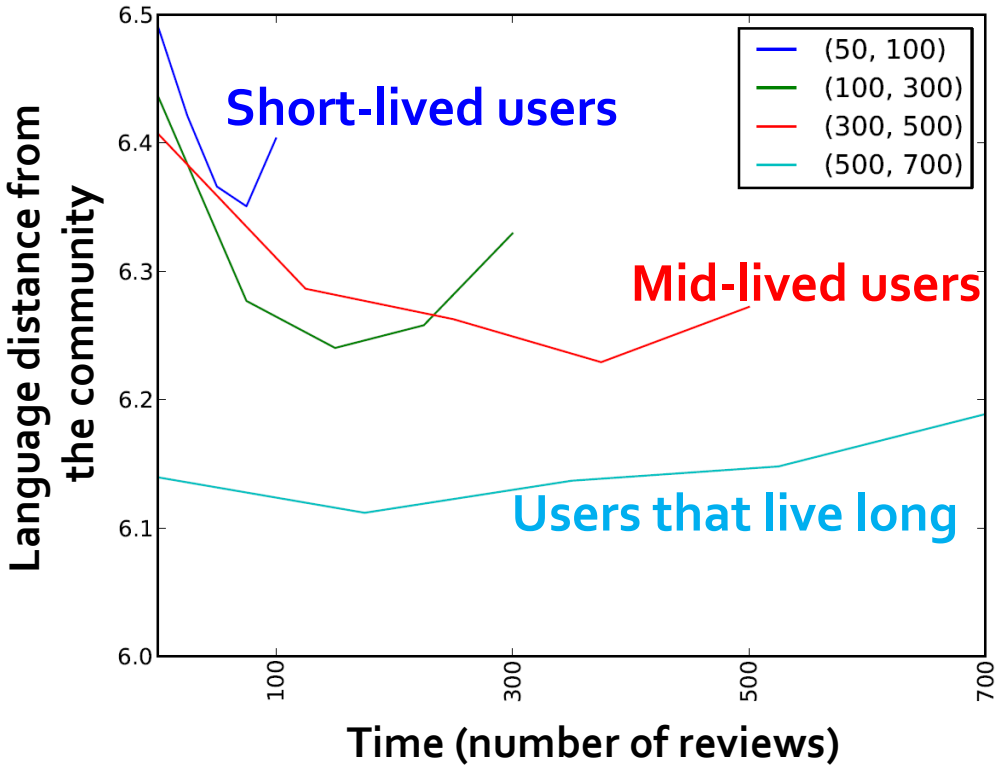
User-Community Change



User-Community Change



All Users Die Old



Regardless of user absolute lifetime, it appears all users go through same life-cycle

Predicting user departures

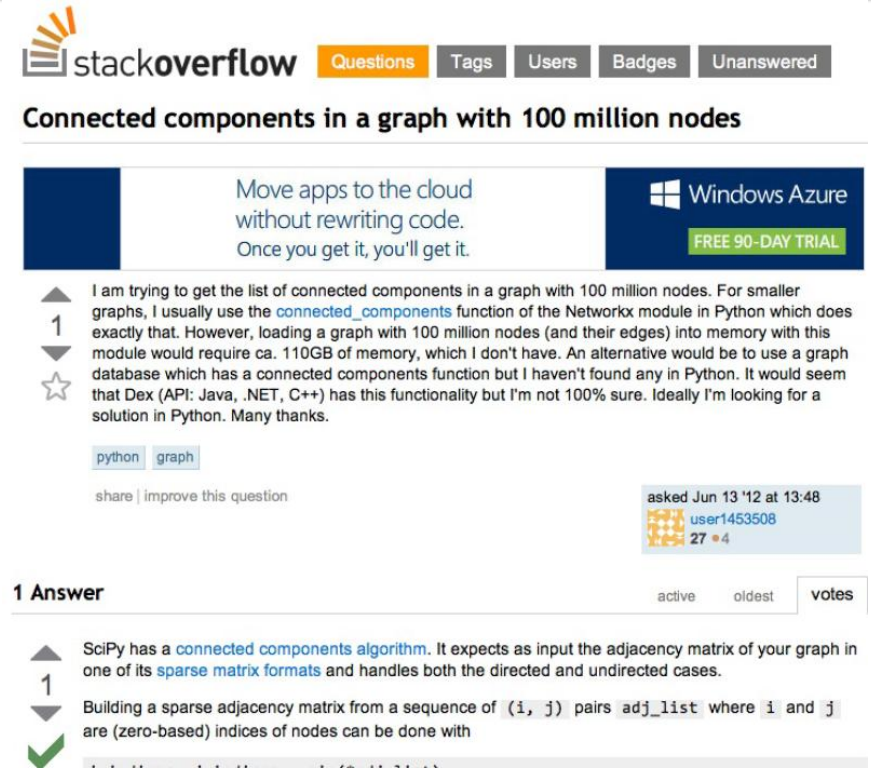
Task: Given user's first 10 reviews. Predict when the user will drop out.

Performance:

Feature	F1
Activity (baseline)	30.5
Language distance	37.4
Adoption of lexical innovations	40.9
First-person singular pronouns	41.2
Review length	42.9

How to motivate activity and contribution?

Stack Overflow



Badges:

Civic Duty badge: Vote at least 300 times

Electorate badge: Vote on at least 600 questions

Newbie: Congrats on your 1st answer

Superstar: You answered 10 questions

Badges

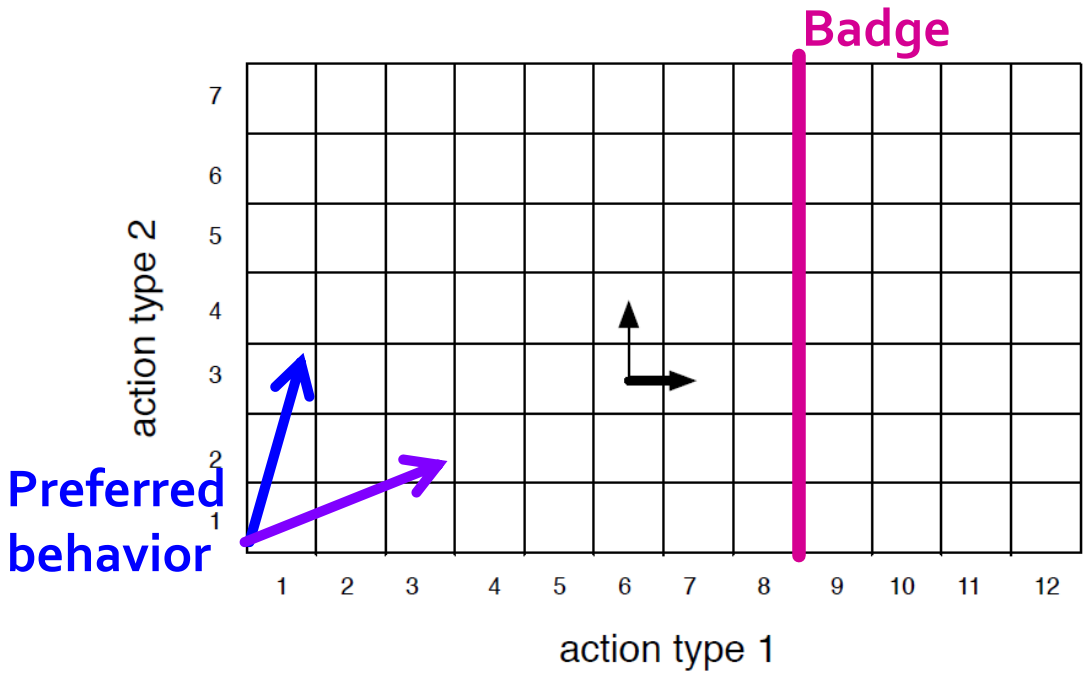
Multiple roles of badges:

- Can recognize a wide range of activities:
  - Total effort, Single high-impact contribution, …
- Serve both as a credential and an incentive
- Gamification

Even though simple, badges are complex:

- How do criteria for a badge translate into effects on user behavior?
- How should site designers dene badges if they want particular outcomes?

Model of Badges



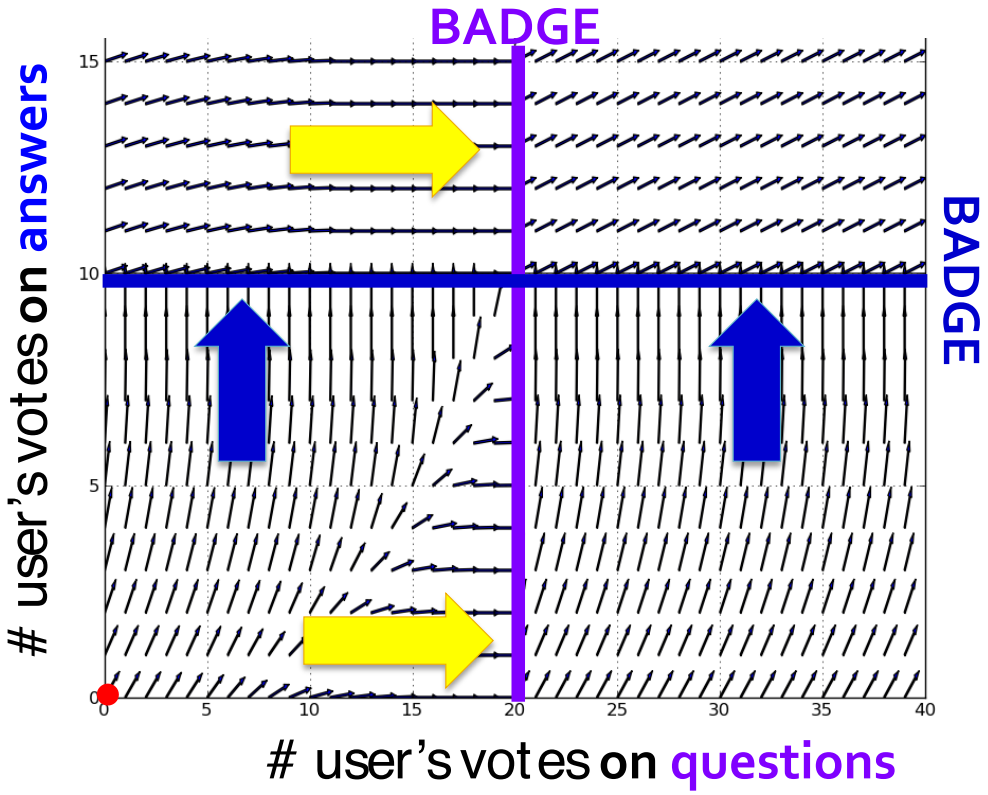
2 parts to the model:

User gains value from obtaining a badge

But it "hurts" user to change behavior

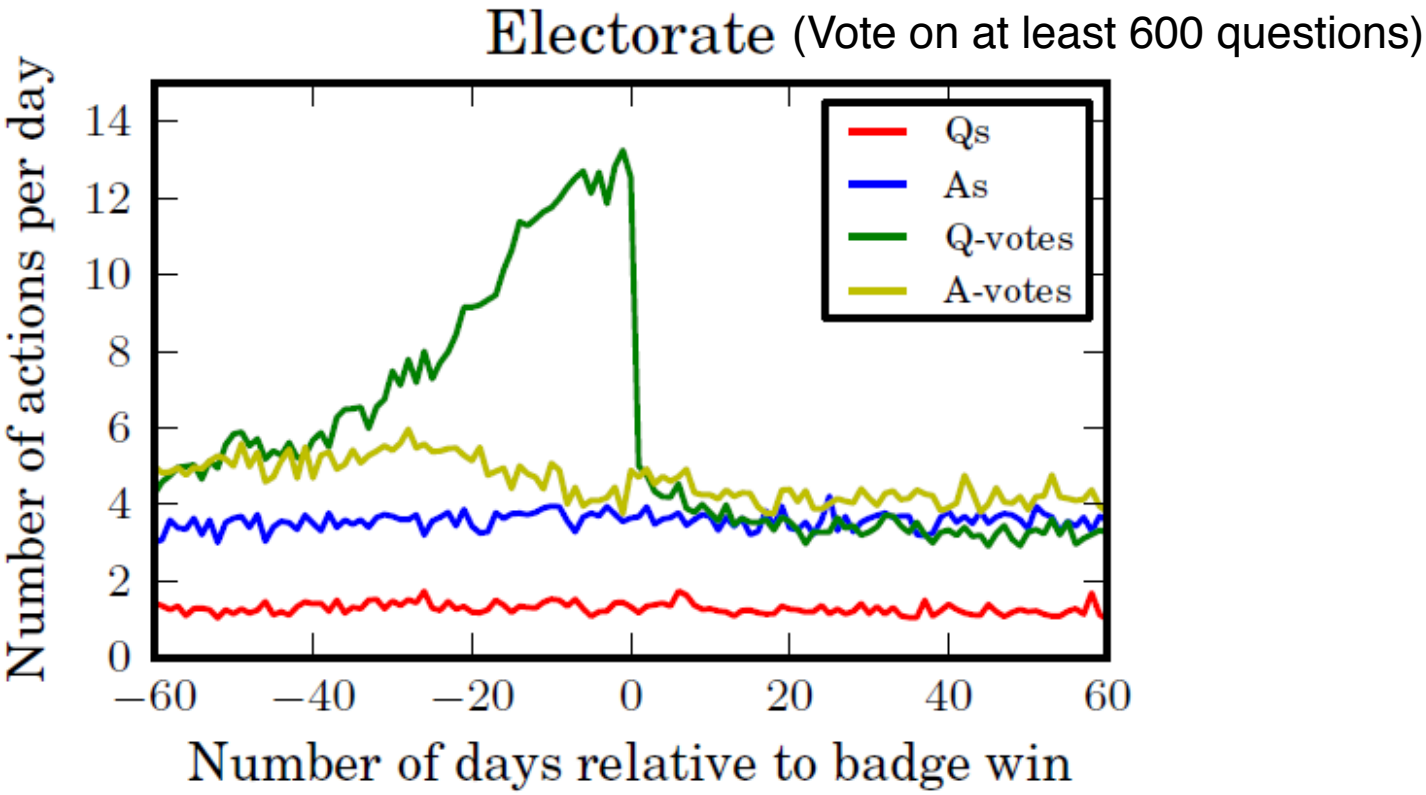
Model Predictions

Influencing user behavior:



Model vs. Data

Model predicts qualitative behavior

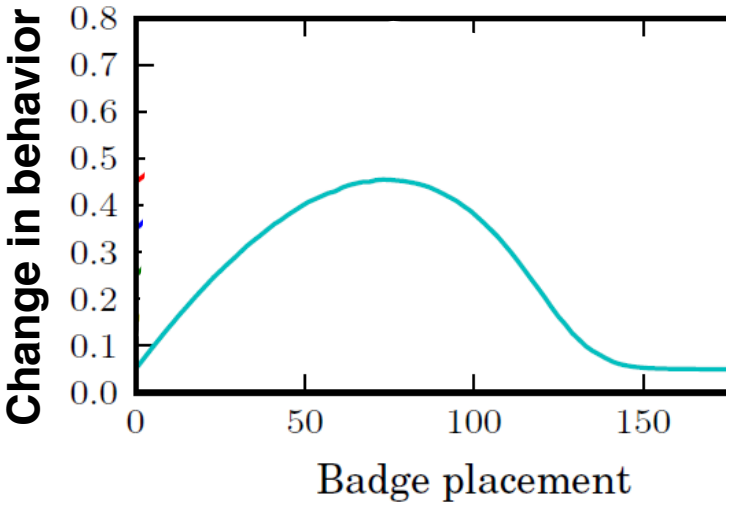


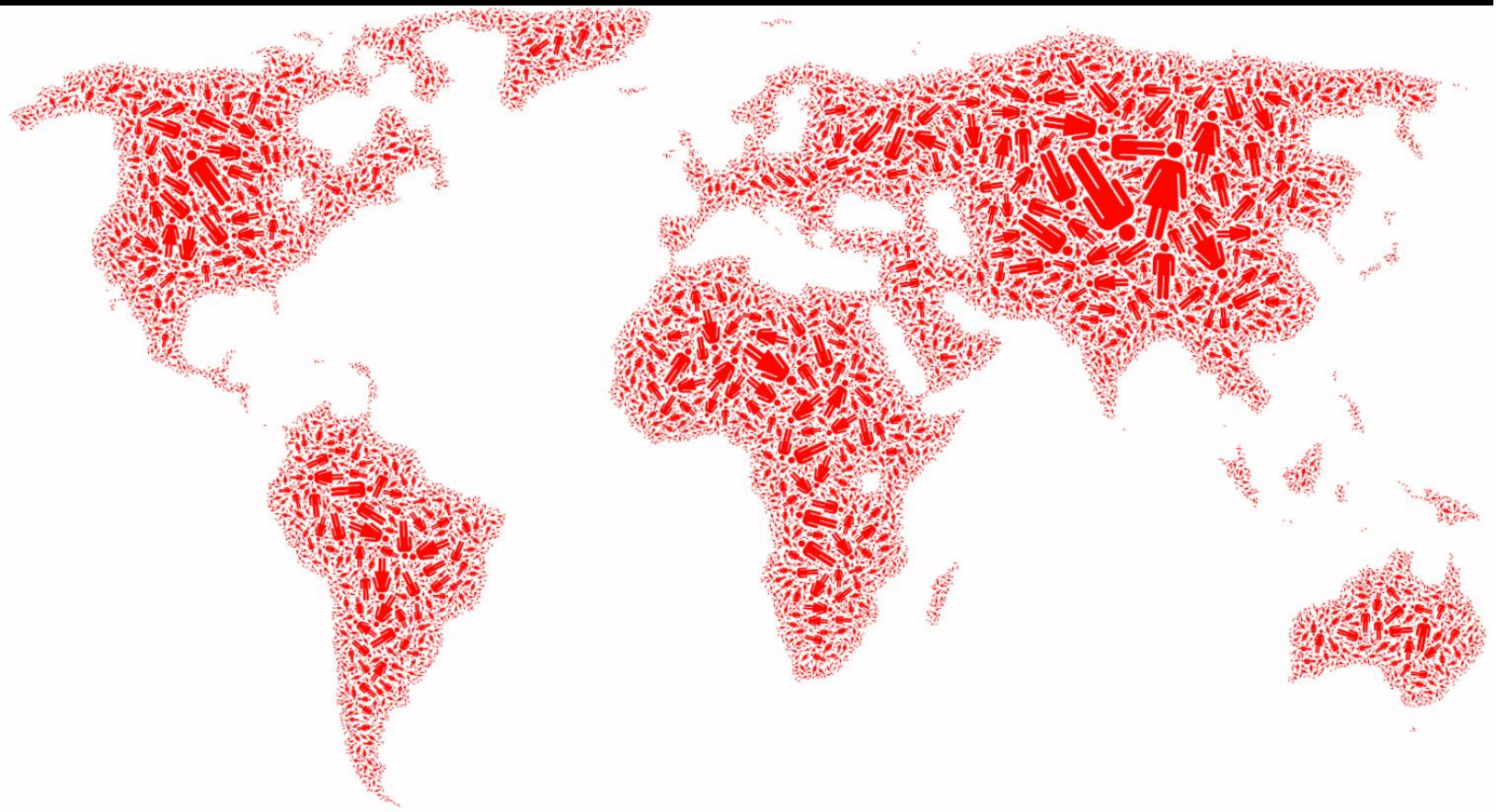
Badge Placement Problem

Question: How should you "place" badges to achieve desired effects?

Our model allows for optimizing the badge placement

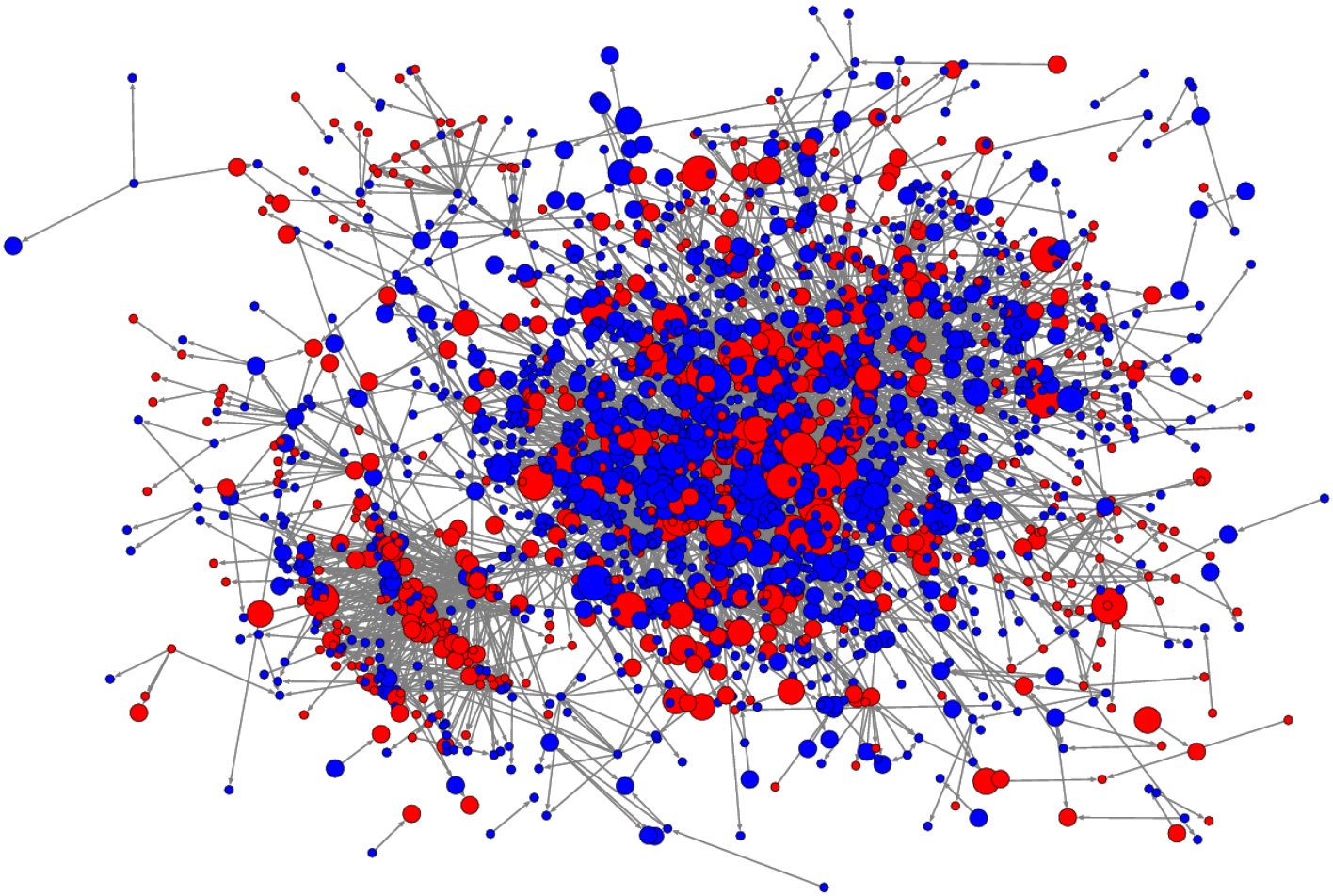
Discussions with Coursera and Folding@home about badge placement for user retention





Web: A sensor into humanity

Networks

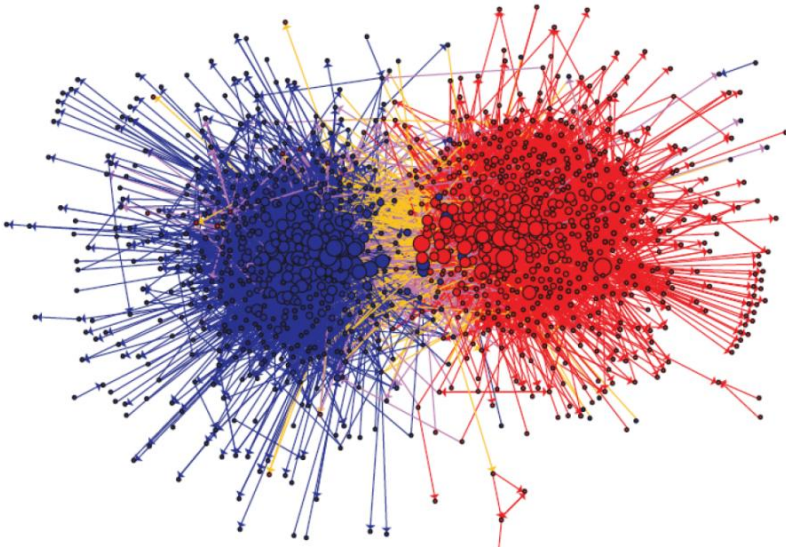


Algorithms for Graphs

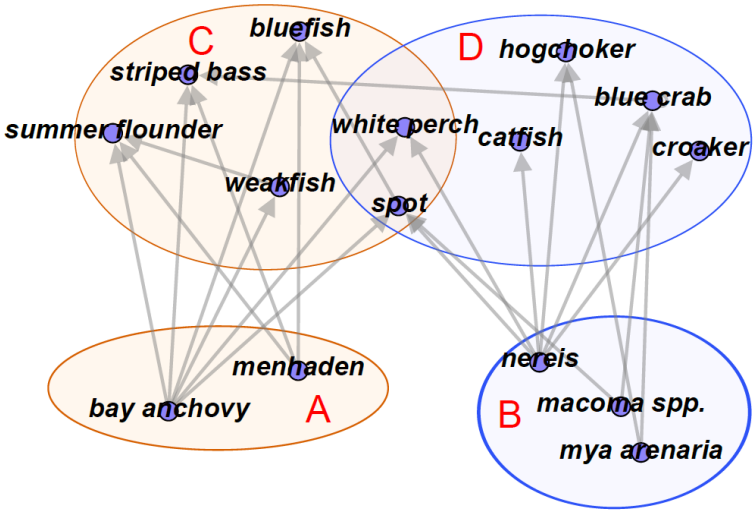
Problem: Finding clusters

Overlapping vs. Non-overlapping

Cohesive vs. Bipartite



U.S. Political blogs

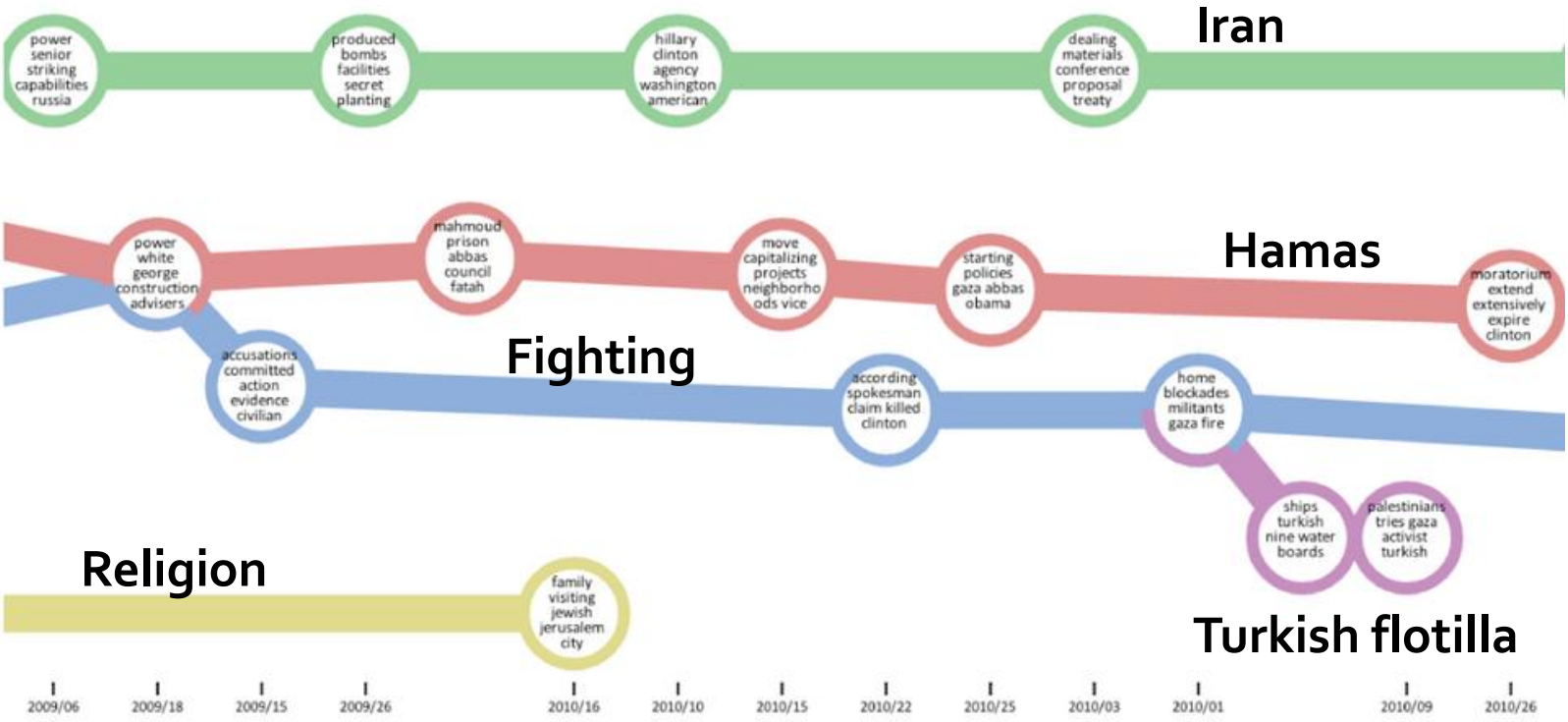


Food web

New Applications

Goal: Want to understand a complex topic

MetroMap of Middle east



Conclusion

What are fundamental patterns of human behavior?

Only recently have basic properties been observed on a large scale

- Confirms social science intuitions; calls others into question

What are good tractable network models?

- Builds intuition and understanding

Benefits of working with large data

- Observe structures not visible at smaller scales


THANKS!
Data + Code:
http://snap.stanford.edu
@jure
Eve Online: Exodus developer CCP publisher CCP
© GameWallpapers.com hosted by JTL.net.com

References

Friendship and Mobility: User Movement In Location-Based Social Networks. E. Cho, S. A. Myers, J. Leskovec. KDD '11.
http://cs.stanford.edu/~jure/pubs/mobile-kdd11.pdf

Steering User Behavior With Badges. A. Anderson, D. Huttenlocher, J. Kleinberg, J. Leskovec. WWW '13.
http://cs.stanford.edu/~jure/pubs/badges-www13.pdf

No Country for Old Members: User lifecycle and linguistic change in online communities. C. Danescu-Niculescu-Mizil, R. West, D. Jurafsky, J. Leskovec, C. Potts. WWW '13.
http://cs.stanford.edu/~jure/pubs/language-www13.pdf