Online social interactions:
a lens on humans
and
a world for humans

Chenhao Tan
chenhao@chenhaot.com
https://chenhaot.com

Teaser trailer: which was shared more on Twitter?

cactus_music
@cactus_music ✔

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition bit.ly/P6GYCq

cactus_music
@cactus_music

I know at some point you've have been saved from hunger by our rolling food trucks friends. Let's help support them! bit.ly/P6GYCq

Only recently have many people had the capability to share messages that reach a wide audience

Power of human traces
on the Internet 15 years ago

The WWW Virtual Library

static hyperlinks
no visible human traces

...mation Management
...ation Quality, Knowledge Management...

...national Affairs
...tional Security, Sustainable Development, UN...

Computing, E-Commerce, Languages, Web...

Arbitration, Law, Environmental Law...

- **Communications and Media**
Communications, Telecommunications, Journalism...

- **Recreation**
Recreation and Games, Gardening, Sport...

- **Education**
Education, Cognitive Science, Libraries, Linguistics...

- **Regional Studies**
African, Asian, Latin American, West European...

- **Engineering**
Civil, Chemical, Electrical, Mechanical, Software...

- **Science**
Biosciences, Health, Earth Science, Physics, Chemistry...

- **Humanities**
Anthropology, Art, History, Museums, Philosophy...

- **Society**
Political Science, Religion, Social Sciences...

http://vlib.org/

3

Power of human traces on the Internet 15 years ago

The WWW Virtual Library

the hyperlinks came from people

...on Management
...uality, Knowledge Management...

...nal Affairs
...Security, Sustainable Development, UN...

...aw, Environmental Law...

...d Games, Gardening, Sport...

- **Education**
  Education, Cognitive Science, Libraries, Linguistics...

- **Engineering**
  Civil, Chemical, Electrical, Mechanical, Software...

- **Humanities**
  Anthropology, Art, History, Museums, Philosophy...

http://vlib.org/

Pagerank [Brin & Page, 1998]
HITS [Kleinberg, 1999]

Google

Much richer social interactions

WIKIPEDIA
The Free Encyclopedia

2001   2004   2006   2008   now

Social Media Landscape 2015

This impacts offline life

change.org

Demand Christian pastor's release from North Korean prison
by Free Pastor Lim · 92,896 supporters · Sign

Urge Congress to address high costs of life-saving medications
by Frederick Candelaria · 64,875 supporters · Sign

Join Tamir Rice's family in calling for Department Justice investigation
by LaTonya Goldsby · 3,813 supporters · Sign

Help mom fighting to save her son's life
by Tonya Carlone · 169,594 supporters · Sign

VICTORY

Congress extends healthcare for September 11 first responders

📍 United States

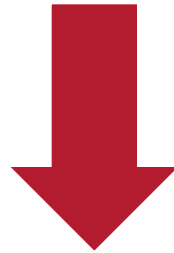👤 186,966 supporters

VICTORY

Taco Bell commits to serving cage-free eggs

📍 United States

👤 166,880 supporters

Online social interactions are an integral part of our lives!

Opportunities & challenges to build better tools/systems for humans

Goal of my research

Micro level
Individual users

Macro level
Service providers

Discover patterns in human behavior and predict human decisions

Improve or build social systems for humans

Micro level
how to change someone else's mind

/r/changemyview

CMV: the Tontine should be legalized and made a common retirement strategy.

[Reference URL omitted] Basically, today we have a huge problem with retirement [...+73 words]

A tontine for retirement looks like [...+56 words] The yearly sum is divided evenly for all the surviving participants [...+25 words].

The key advantages as I see it are:

*We don't need actuaries [...+29 words...]

*Management fees can be quite low [...+22 words]

* [Another reason]

* [Another reason]

But CMV. Are there major risks I am not forseeing? [+2 more questions]

A tontine is a pretty crappy retirement vehicle for most people. It pays out the least when you need the most, and the most when you need the least.

People's income needs in retirement generally fall as they age. [...+35 words]
[URL]

Very interesting. I'll give a Δ because I didn't have any idea that was true and changes my idea of how the tontine should work. That said, I don't think it's unsolvable [...+44 words]
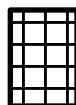
The Social Security system is basically one giant Tontine [...+17 words]

10 more comments

Macro level
how to maintain healthy dynamics

Comment Rules:

hover over sections for more info

1. Direct responses to a CMV post must challenge at least one aspect of OP's stated view (however minor), unless they are asking a clarifying question. ▾

2. Don't be rude or hostile to other users. ▾

3. Refrain from accusing OP or anyone else of being unwilling to change their view. ▾

4. If you have acknowledged/hinted that your view has changed in some way, please award a delta. ▾

5. No low effort comments. ▾

stack**overflow**

Badge system design
[Anderson et al. 2013, Immorlica et al. 2015]

coursera edX

User retention (single community)
[Danescu-Niculescu-Mizil et al. 2013, Dasgupta et al. 2008, Dror et al. 2012, Rowe 2013, Yang et al. 2013, etc]

Community rules
[Kraut et al. 2012, Rheingold 2000]

Overview of my thesis

Micro level: effect of language	Macro level: communities and networks

Overview of my thesis

Micro level: effect of language	Macro level: communities and networks

Message sharing [ACL'14]

Users' multi-community engagement [WWW'15]

Winning arguments [WWW'16]

/r/atheism
/r/trueatheism

Highly related communities [ICWSM'16]

polymath Internet collaboration [WWW'16]
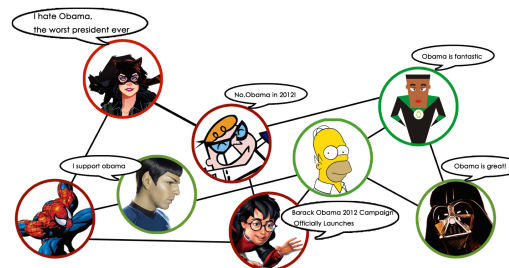
arXiv.org Paper revision [ACL'14 short]
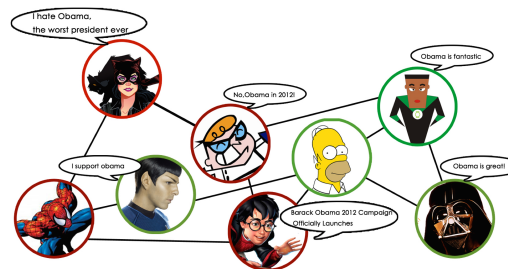
Network & topics /behavior/emotion [ICWSM'13, KDD'11, KDD'10]

Understand/predict human decisions
Improve social systems for humans

Message sharing [ACL'14]

New approaches to classic problems

Winning arguments [WWW'16]

Network & topics/behavior/emotion [ICWSM'13, KDD'11, KDD'10]

polymath Internet collaboration [WWW'16]

/r/atheism
/r/trueatheism

Highly related communities ['16]

New problems from new behavioral data

arXiv.org Paper revision [ACL'14 short]

reddit

Users' multi-community engagement [WWW'15]

Overview of my thesis

Micro level: effect of language	Macro level: communities and networks

Message sharing [ACL'14]

Users' multi-community engagement [WWW'15]

Winning arguments [WWW'16]

/r/atheism
/r/trueatheism
Highly related communities [ICWSM'16]

polymath
Internet collaboration [WWW'16]

arXiv.org
Paper revision [ACL'14 short]

Network & topics /behavior/emotion [ICWSM'13, KDD'11, KDD'10]

Overview of my thesis

Micro level: effect of language	Macro level: communities and networks

 Message sharing [ACL'14]

 Users' multi-community engagement [WWW'15]

 Winning arguments [WWW'16]

/r/atheism /r/trueatheism   Highly related communities [ICWSM'16]

polymath   Internet collaboration [WWW'16]

arXiv.org   Paper revision [ACL'14 short]

 Network & topics /behavior/emotion [ICWSM'13, KDD'11, KDD'10]

Getting more people to share on Twitter

cactus_music
@cactus_music

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition bit.ly/P6GYCq

Important factors for sharing

author status, topic popularity
[Milkman and Berger, 2012; Romero et al. 2013; Suh et al. 2010; etc]

cactus_music
@cactus_music

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition
bit.ly/P6GYCq

Important factors for sharing

author status, topic popularity
[Milkman and Berger, 2012; Romero et al. 2013; Suh et al. 2010; etc]

cactus_music
@cactus_music

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition bit.ly/P6GYCq

Barack Obama ✓
@BarackObama

✔

Four more years.
pic.twitter.com/bAJE6vom

Wording is entirely under the author's control

cactus_music
@cactus_music

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition bit.ly/P6GYCq

cactus_music
@cactus_music

I know at some point you've have been saved from hunger by our rolling food trucks friends. Let's help support them! bit.ly/P6GYCq

It is all about followers (Score:3, Interesting)

by mysterons (1472839) on Thursday May 15, 2014 @01:36PM (#47010441)

We did a study on predicting when a tweet would be retweeted (this paper cites us). The dominant factor is not what you write, but how many followers you have. Basically, a famous person can write anything and it will be retweeted. An unknown person can write the same tweet and it will be ignored.

Link to paper:

Sasa Petrovic, Miles Osborne and Victor Lavrenko. RT to win! Predicting Message Propagation in Twitter. ICWSM, Barcelona, Spain. July 2011. http://homepages.inf.ed.ac.uk/...[ed.ac.uk]

Reply to This   Share

Lessons from science: experiments

"How important is it to you to be a voter in the upcoming election?" ✓

"How important is it to you to vote in the upcoming election?"
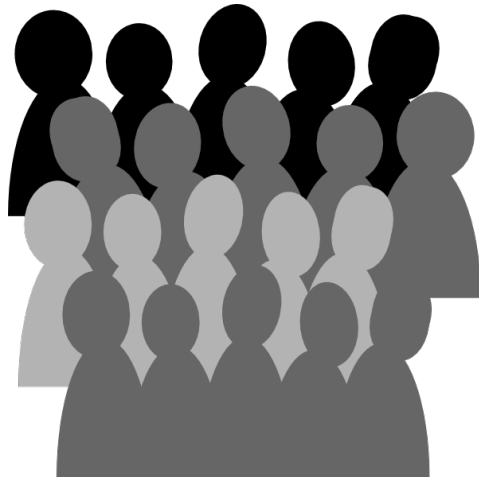
Representative group A

Representative group B

Bryan, Walton, Rogers and Dweck 2011
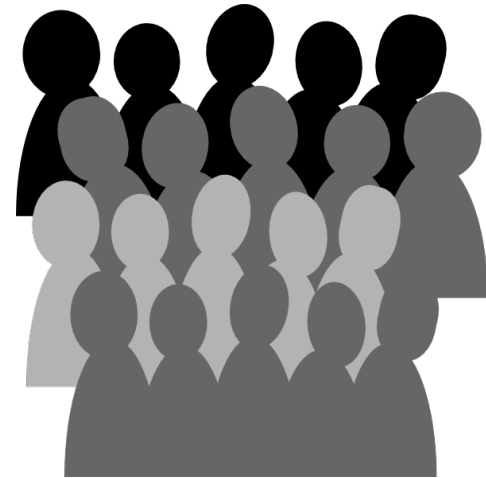
Experiments are great, but they are difficult to scale

- Requires recruiting participants and asks for extra effort from participants
- Requires experiment designers to propose different wordings
- Lab can be different from real life

Natural Experiment Paradigm

- *Same* speaker
- conveying the *same* info
- *Same* situation
- Varies their wording

and see the effects

http://www.imdb.com/title/tt0107048/
More: The Butterfly Effect, Edge of Tomorrow

The same users post multiple tweets on the same topic

Topic- and author-controlled pairs

cactus_music
@cactus_music

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition bit.ly/P6GYCq

cactus_music
@cactus_music

I know at some point you've have been saved from hunger by our rolling food trucks friends. Let's help support them! bit.ly/P6GYCq

The same users post multiple tweets on the same topic

GeorgeMonbiot
@GeorgeMonbiot

read @ameliagentailman's report today, then tell me Tories are no longer the nasty party: guardian.co.uk/society/2012/o…

GeorgeMonbiot
@GeorgeMonbiot ✔

Work capability tests: designed by bastards, performed by idiots. guardian.co.uk/society/2012/o…

https://chenhaot.com/retweetedmore/quiz
http://www.nytimes.com/interactive/2014/07/01/upshot/twitter-quiz.html

Topic- and author-controlled pairs are actually common!

- *2.4 million* topic- and author-controlled tweet pairs after processing the complete history of millions of users

- Careful filtering to reach a situation that is close to natural experiments (11K pairs)
  - \>5K followers, <12 hours
  - Non-trivial textual changes
  - Significant changes in retweet numbers
  - Limit the number of pairs by an author to 50

Does wording matter?

Wording does not matter

↓

Humans/machines should not be able to tell which one in a pair was retweeted more

Humans/machines can tell which one in a pair was retweeted more (accuracy > 50%)

↓

Wording matters!

Can humans tell which tweet will be retweeted more?

- Randomly sample 100 pairs
- 20 pairs a task on Amazon Mechanical Turk
- 39 judgments for each pair

Can humans tell which tweet will be retweeted more?
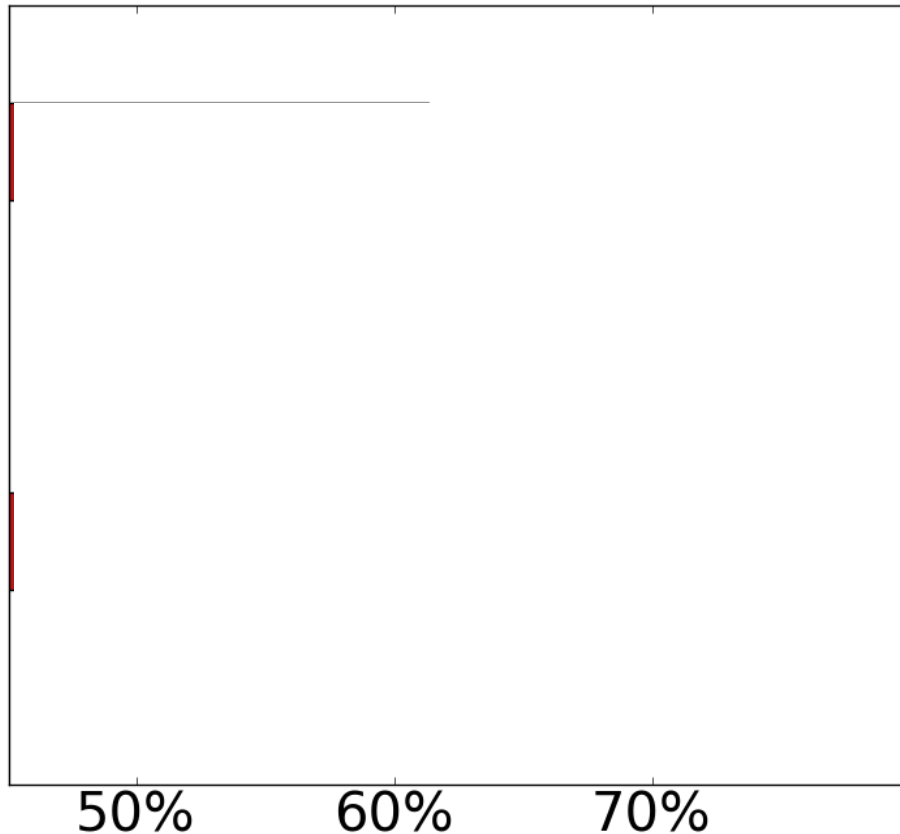
Average accuracy for each labeler: 61.3%

Accuracy of the majority label for each pair: 73%

Predict which tweet will be retweeted more within a pair

- Features
  - Custom features that we proposed: informativeness, etc (39 features)
  - Bag of words: unigram+bigram (7K features)
- Approach
  - Take the difference between features for two tweets in a pair
  - Logistic regression

Predict which tweet will be retweeted more within a pair

- 5-fold *cross validation* experiments

1	2	3	4	5

- *Heldout* experiments

fresh

Only used once, 6 days before submission!

Predict which tweet will be retweeted more within a pair

- A strong baseline

a classifier that uses tweets without any control for training

Most retweeted	Least retweeted

Use bag-of-words features, [number of followers and timing]

Cross validation accuracy 98.8%

Cross-validation performance: is control necessary?

Unpaired bag-of-words

Average human accuracy (on a sample of 100 pairs)
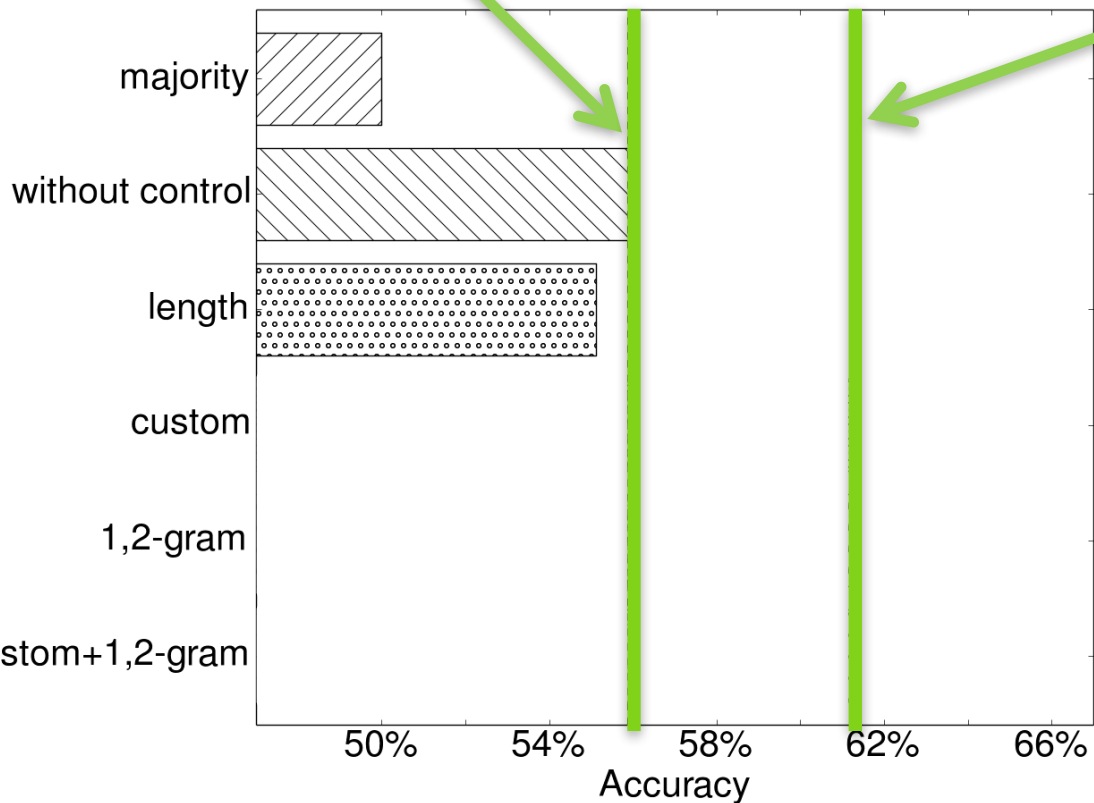
- More than 10% absolute improvement over the baseline
- Potentially interesting implication for machine learning

Cross-validation performance

Unpaired bag-of-words
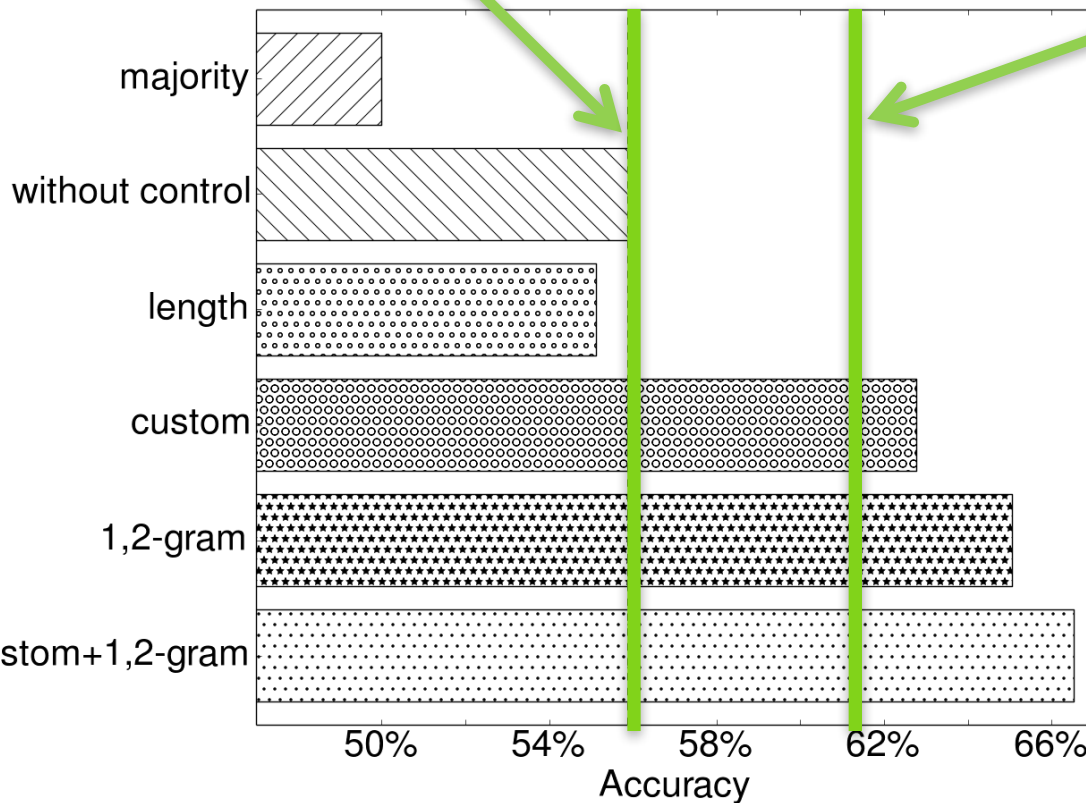
Average human accuracy (on a sample of 100 pairs)

- More than 10% absolute improvement over the baseline
- Custom does pretty well by itself, and outperforms average human accuracy
- Adding custom improves bag-of-words

Same results hold in heldout data

Unpaired bag-of-words
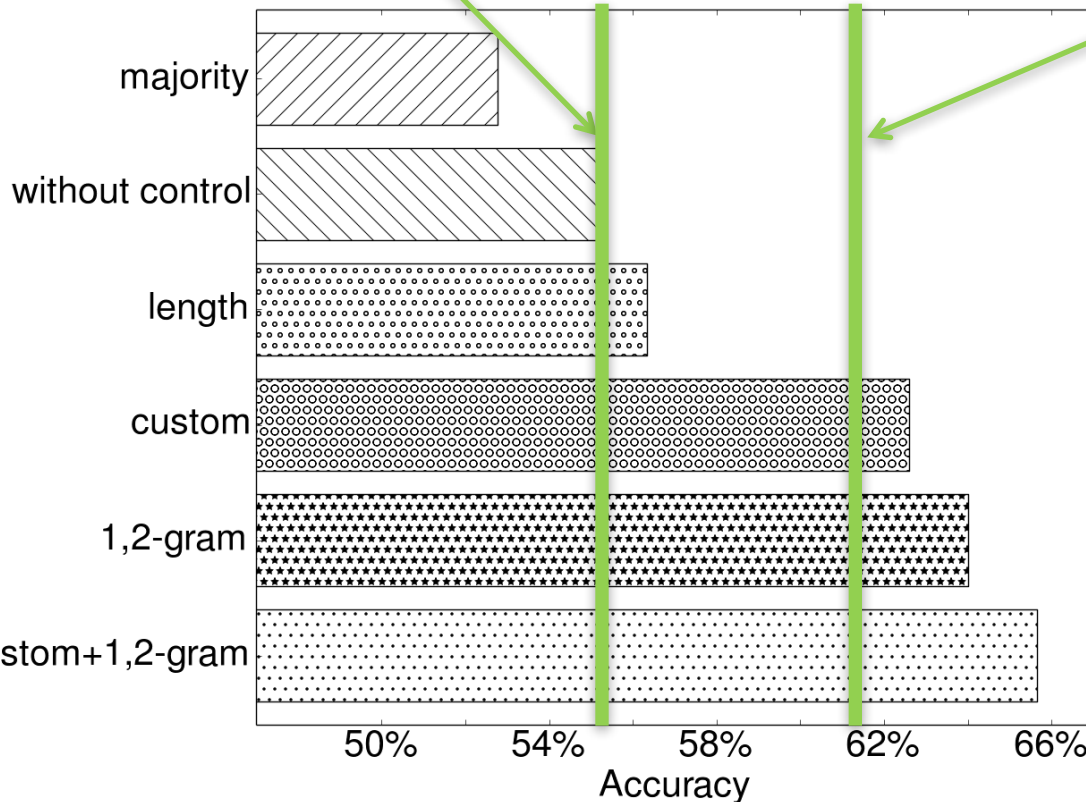
Average human accuracy (on a sample of 100 pairs)

- More than 10% absolute improvement over the baseline
- Custom does pretty well by itself, and outperforms average human accuracy
- Adding custom improves bag-of-words

https://chenhaot.com/retweetedmore

Should we conform to community norm?

- Memorable movie quotes corresponds to lower language similarity [Danescu-Niculescu-Mizil et al. 2012]

Should we conform to community norm?

Compute unigram, bigram language model score for a tweet according to Twitter (unigram as an example):

$$\frac{1}{|t|} \sum_{w \in t} log_2 P(w)$$

higher score = more similar to twitter language

Average log probability for a tweet

Probability according to a large (non-paired) tweet sample

Be like the community (conformity)

unigram, bigram language model score according to Twitter

higher score = more similar to twitter language

	Effective?
Twitter unigram similarity	↑ ($p < 0.001$)
Twitter bigram similarity	↑ ($p < 0.001$)

Should we maintain personal style?

Compute unigram, bigram language model score for a tweet according to personal history (unigram as an example):

$$\frac{1}{|t|} \sum_{w \in t} log_2 P(w)$$

higher score = more similar to personal history

Average log probability for a tweet

Probability according to a user's own history

Be true to yourself

unigram, bigram language model score according to personal history

higher score = more similar to personal history

	Effective?
Personal unigram similarity	⬆ (p < 0.001)
Personal bigram similarity	———

Micro level summary

- Natural experiments based on online social interactions show that language matters in message propagation!

- Controlling topics and authors can improve predictive performance significantly over an approach without control

- Be like the community and be true to yourself

 [We also used informativeness, headline similarity, emotions, generality, personal pronouns, generality, readability]

https://chenhaot.com/retweetedmore

Overview of my thesis

Micro level: effect of language	Macro level: communities and networks

Message sharing [ACL'14]

Users' multi-community engagement [WWW'15]

Winning arguments [WWW'16]

/r/atheism /r/trueatheism — Highly related communities [ICWSM'16]

polymath — Internet collaboration [WWW'16]

Network & topics /behavior/emotion [ICWSM'13, KDD'11, KDD'10]

arXiv.org — Paper revision [ACL'14 short]
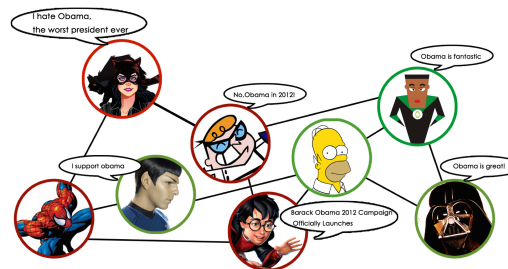
Predict future user activity in a multi-community setting

- An active platform where users submit posts, make comments and rate posts
- Many sub-communities (subreddits)

User trajectories across communities

changemyview

politics

Babybumps

...

How much do users explore new communities?

Food trajectory analogy

#unique communities

Reddit age

How much do users explore new communities?

"Users get old" hypothesis: Biography studies suggest that people settle down after an initial period of exploration. [C. Bühler 1935]

#unique communities
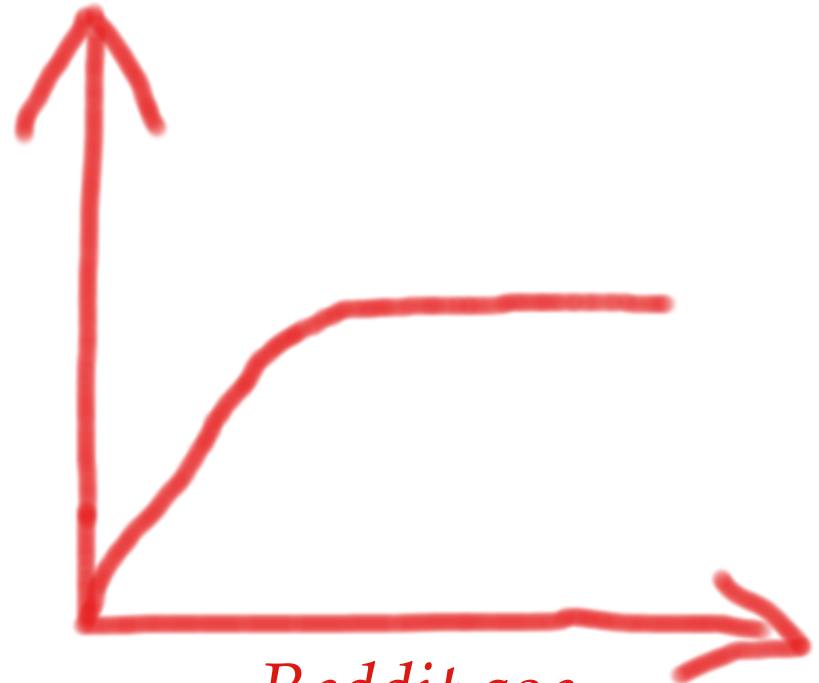
Reddit age

People stop adapting their language in a single-community setting (BeerAdvocate)

[Danescu-Niculescu-Mizil et al. 2013]
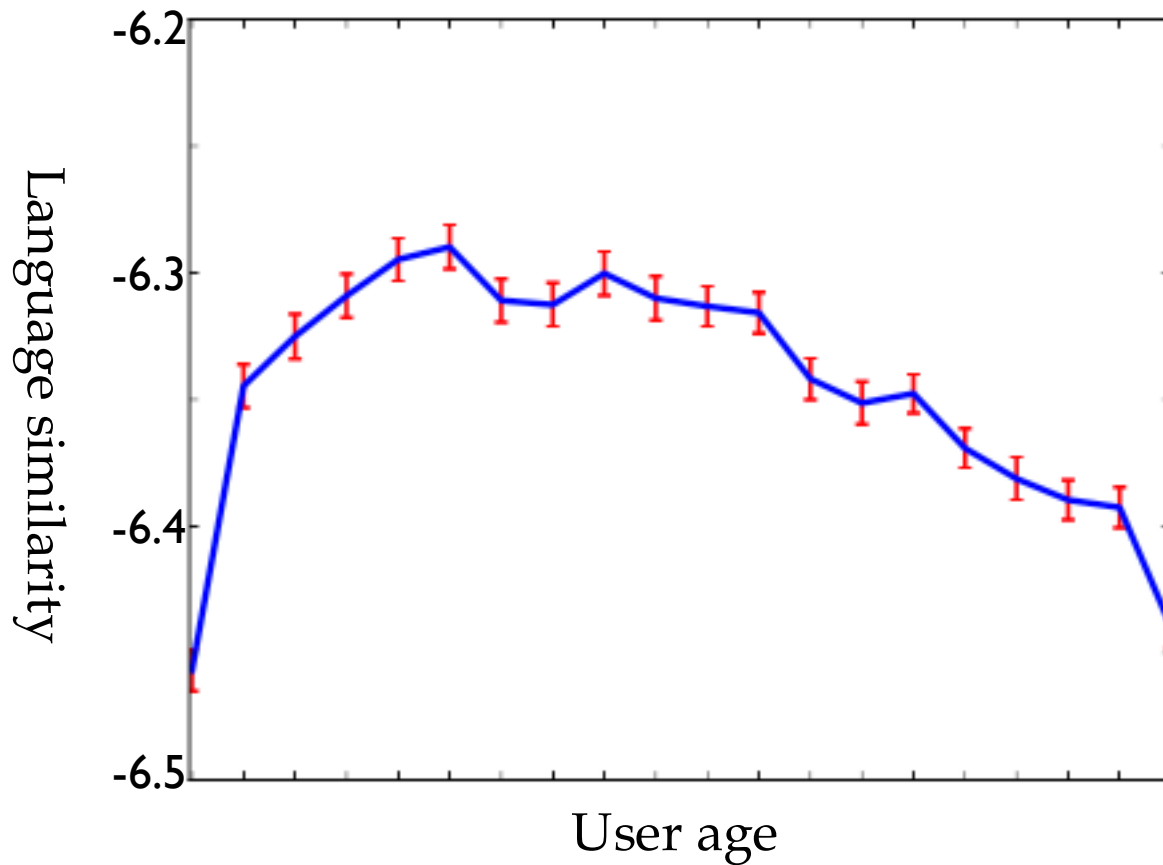
First 50 posts on Reddit

Reddit
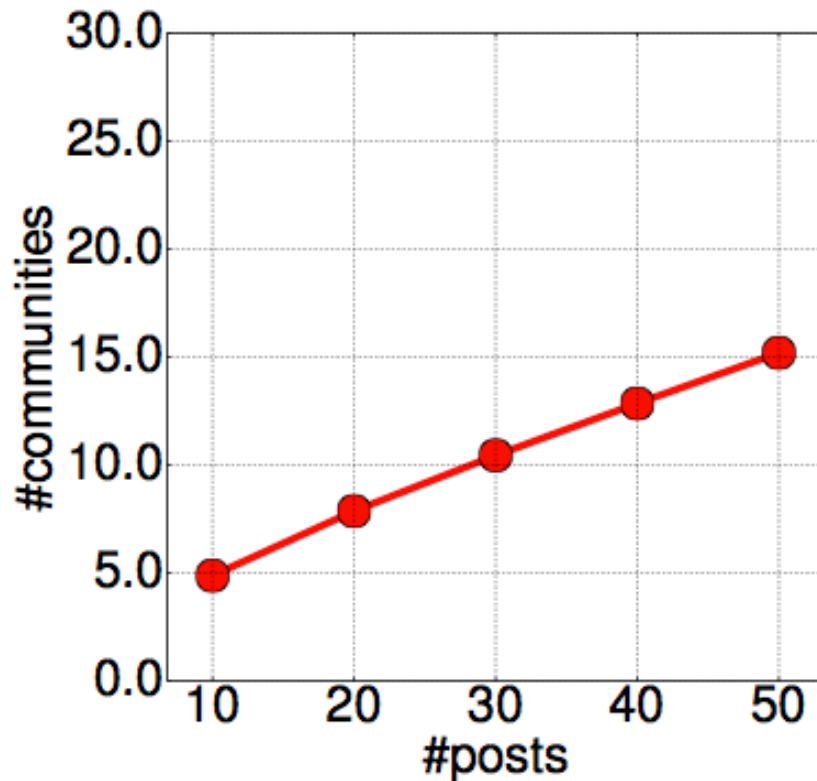
Error bars (tiny) show standard error.

The average time to accumulate 50 contributions is 456.0 days on Reddit.
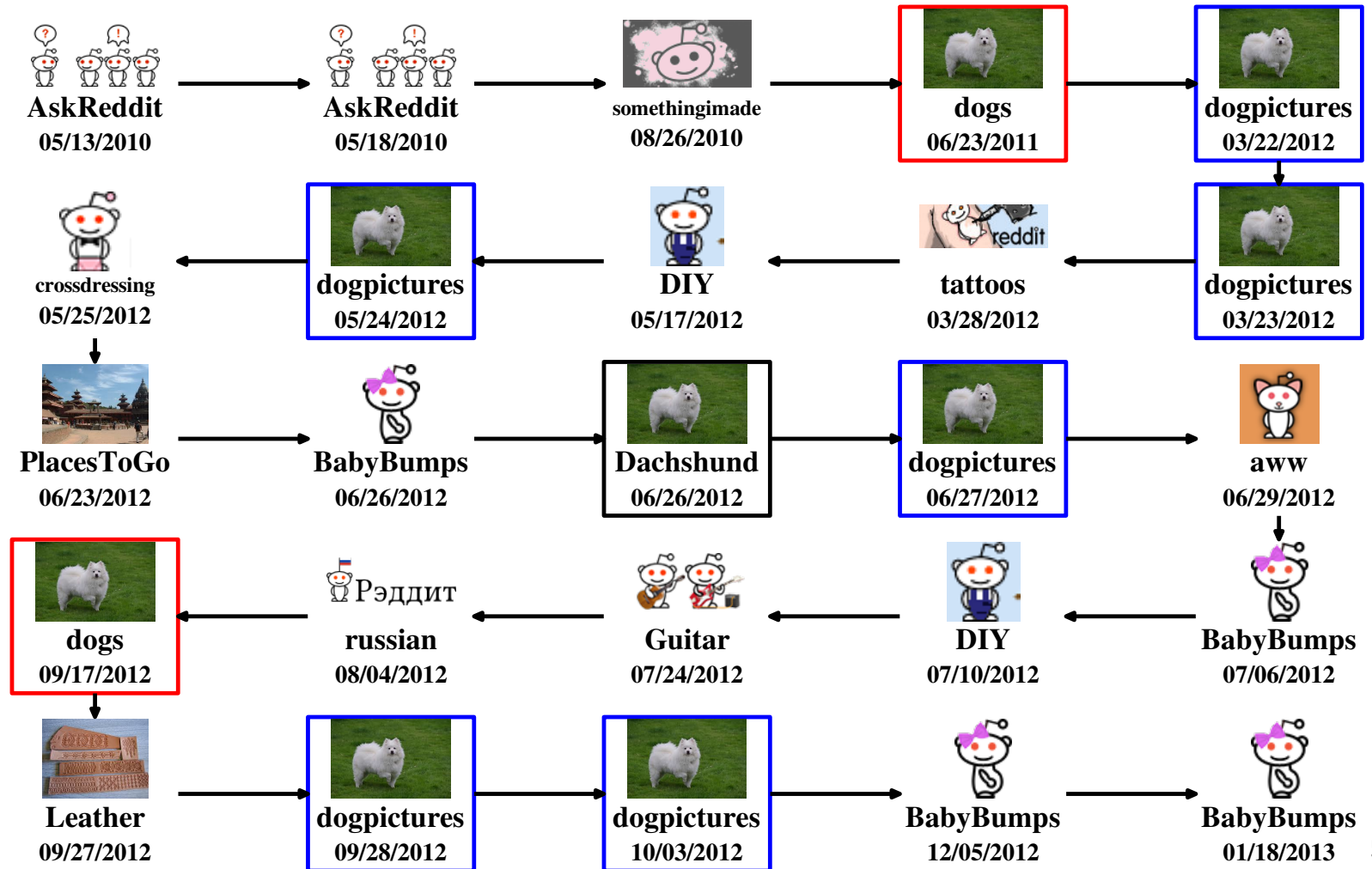
First 50 posts on Reddit and DBLP
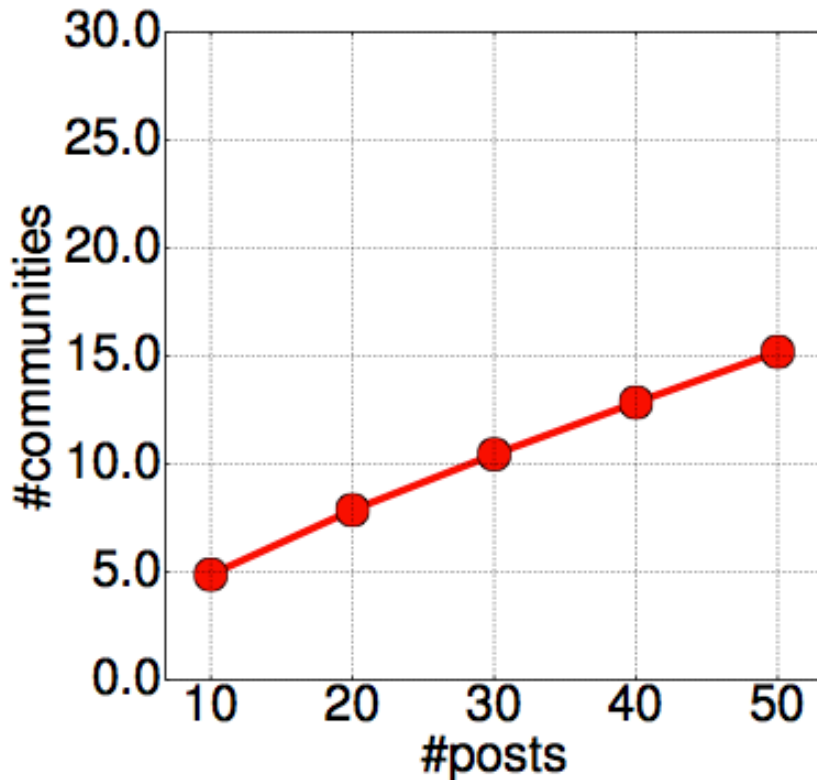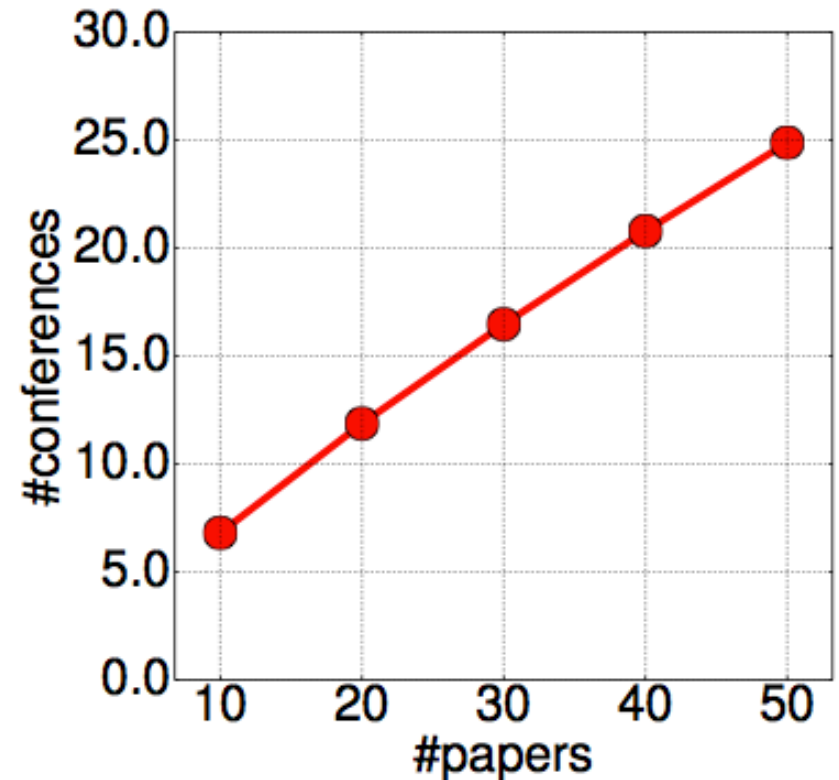
Reddit

DBLP

Error bars (tiny) show standard error.

The average time to accumulate 50 contributions is 456.0 days on Reddit, 15.6 years on DBLP.

Lifetime on Reddit and DBLP
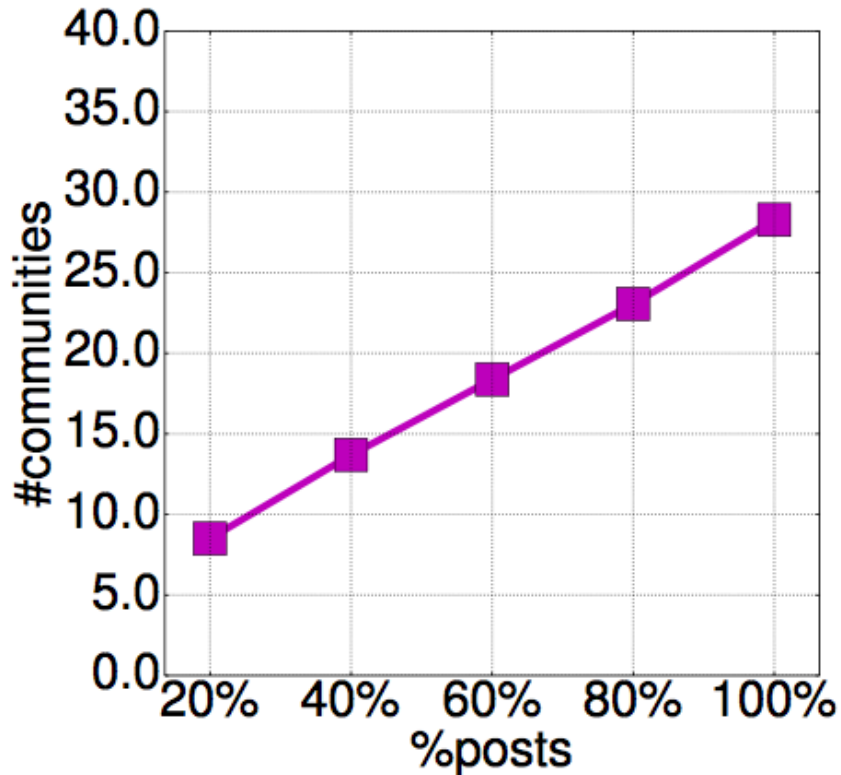
Reddit

DBLP
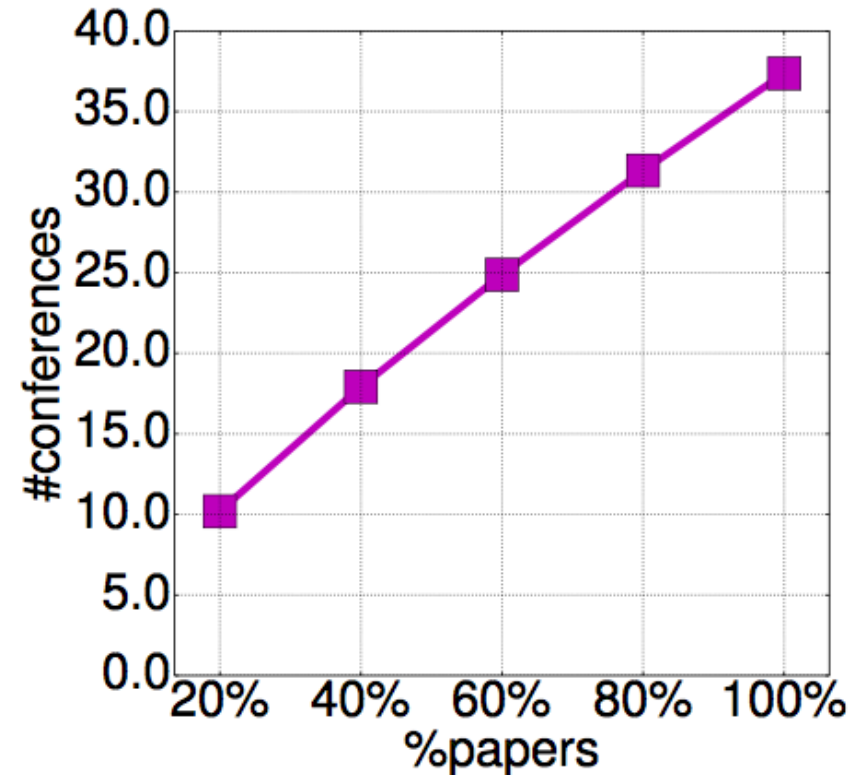
Error bars (tiny) show standard error.

# User trajectories across communities

An example from a user on Reddit

Main dataset: Reddit

- *All 76.6M posts ever submitted to Reddit* from its inception until Jan 2014
- 157K "50+" posters who first posted between Jan 2008 and Jan 2012 [Danescu-Niculescu-Mizil et al. 2013]
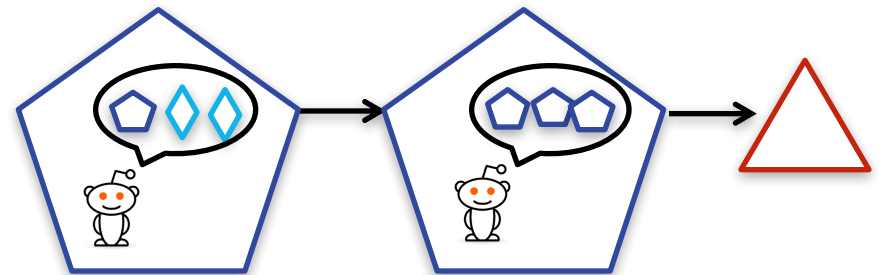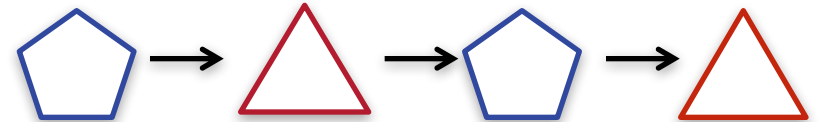
Link: *https://chenhaot.com/pages/multi-community.html*

Predict whether users will abandon Reddit from user trajectories

- (How) does the wandering pattern change over time?

- Do people adapt their language in each community over time?

- Do people receive better evaluations over time?
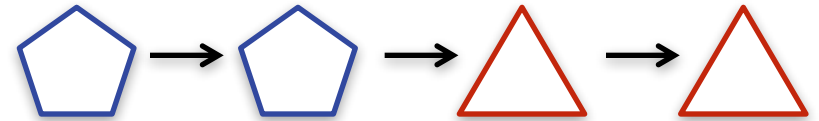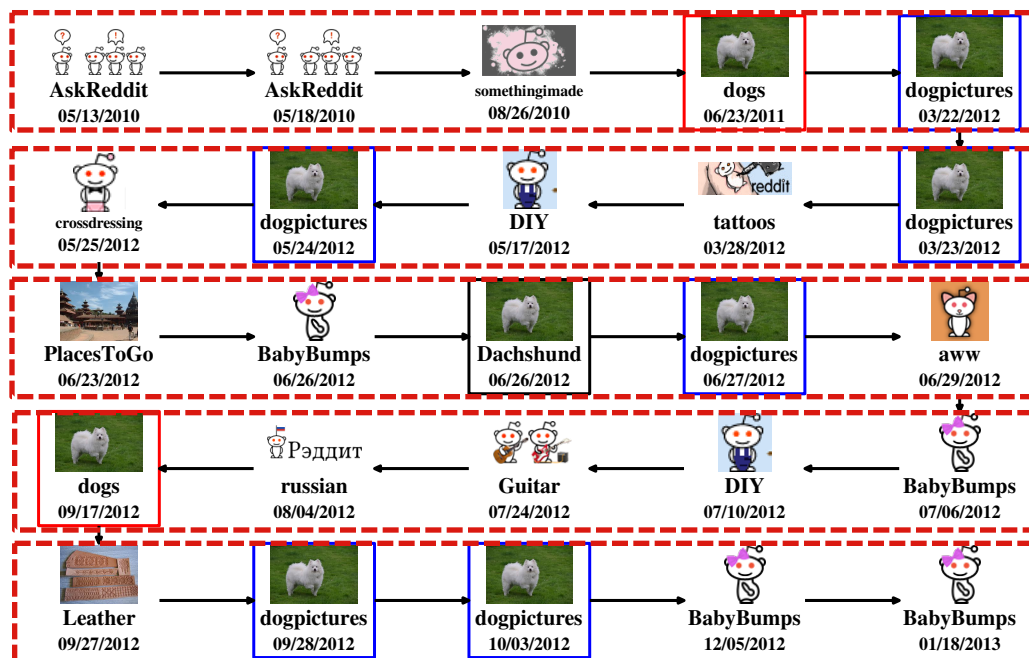

OR?

evaluations=
ups - downs

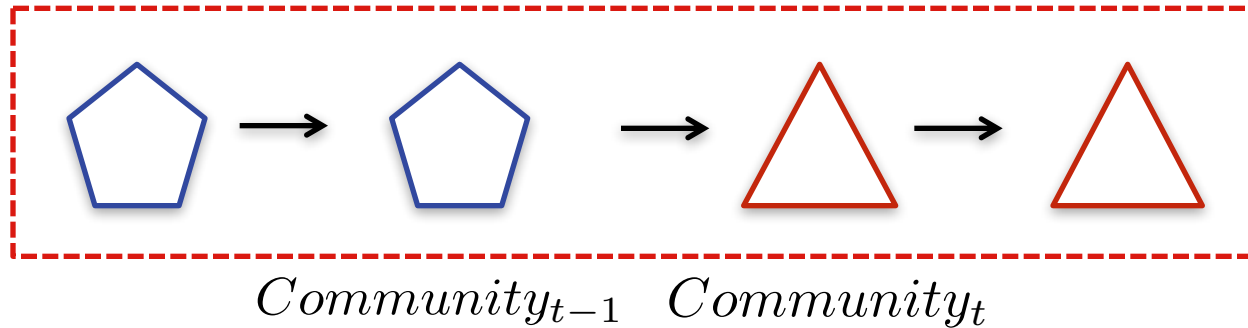# A framework for measuring properties of the trajectory

e.g., window size = 5

Split the trajectory into windows of the same size *(w=10 in main experiments)*

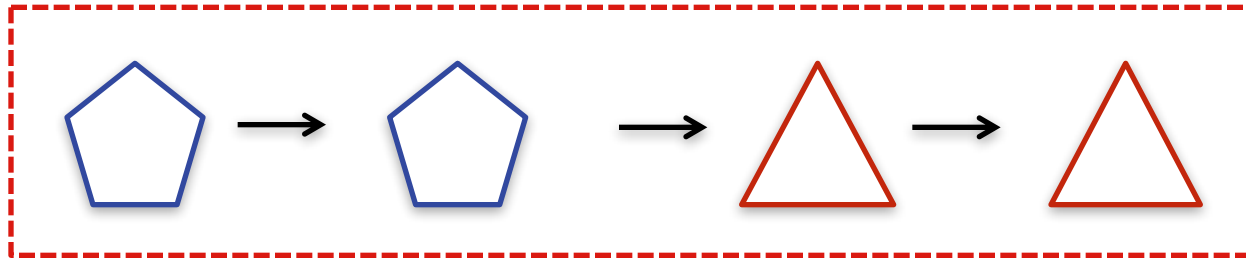Define a function (*F*) on a window to capture different properties and obtain a time series

Users' "jumping" behavior

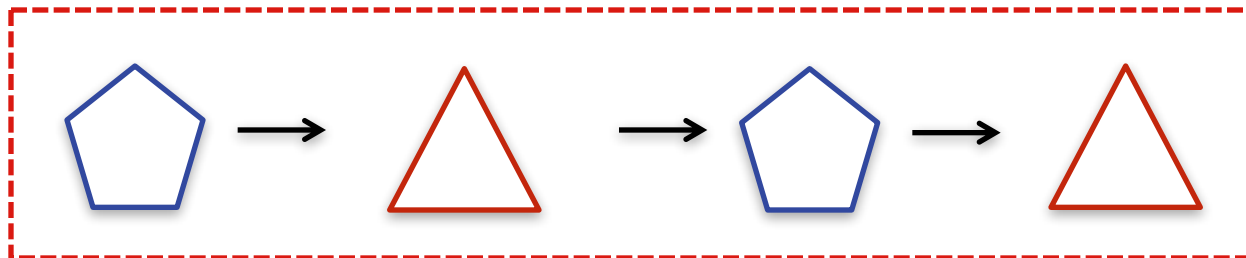F: Count of $Community_t \neq Community_{t-1}$ in a window

$$Community_{t-1} \quad Community_t$$

Users' "jumping" behavior

F: Count of $Community_t \neq Community_{t-1}$ in a window

$$Jumps = 1$$

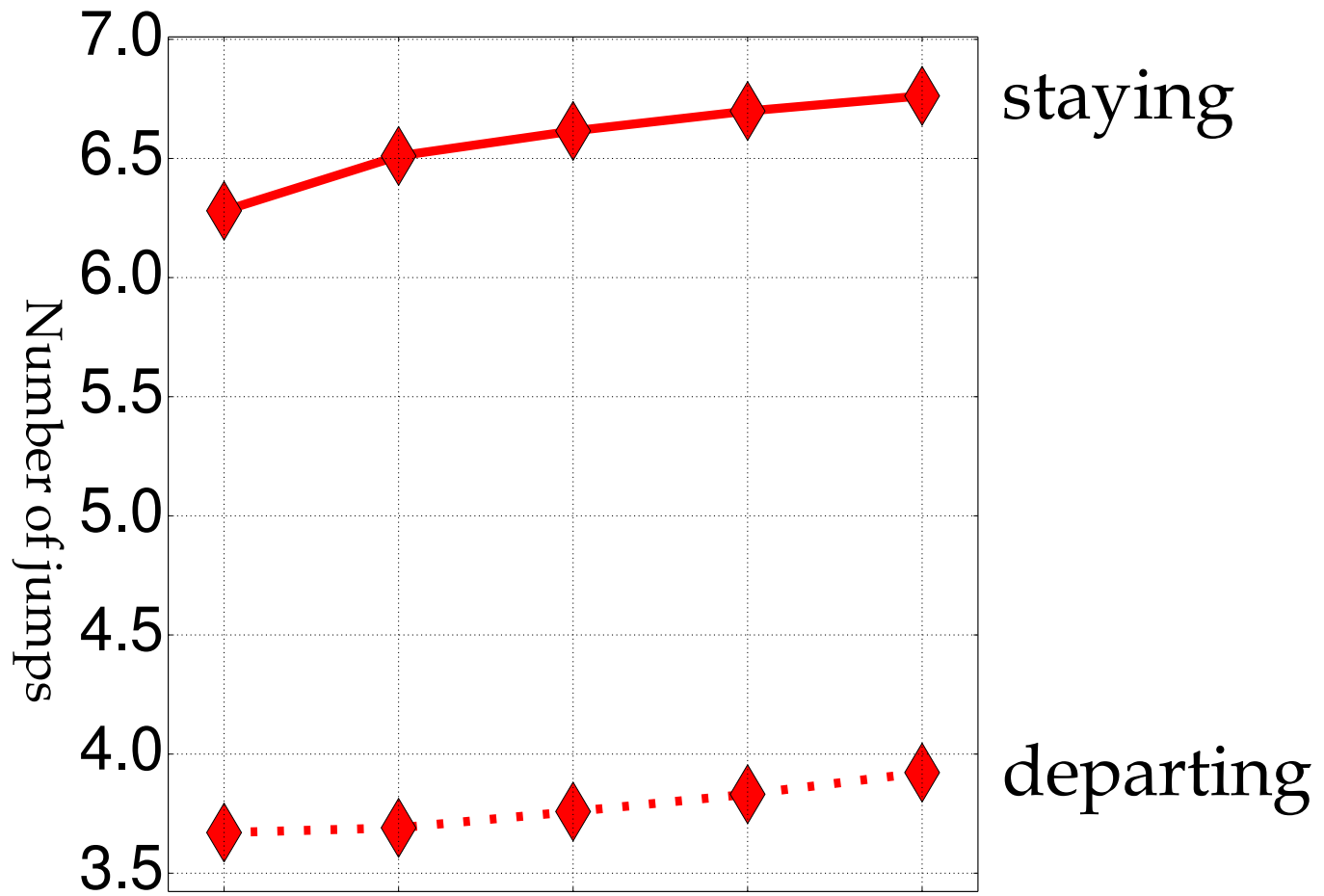$$Jumps = 3$$

Differences between future departing users and future staying users

- Future departing users: stopped posting in the entire Reddit in the last 6 months (44K)

- Future staying users: stay *active* in the last 6 months (76K)

Do departing users or staying users jump more?

Users "jump" more over time; future departing users less so

staying

departing

Error bars (tiny) show standard errors.

Users get more adventurous over time; future departing users less so

Many more different perspectives on the wandering pattern:

number of unique communities, ⬆
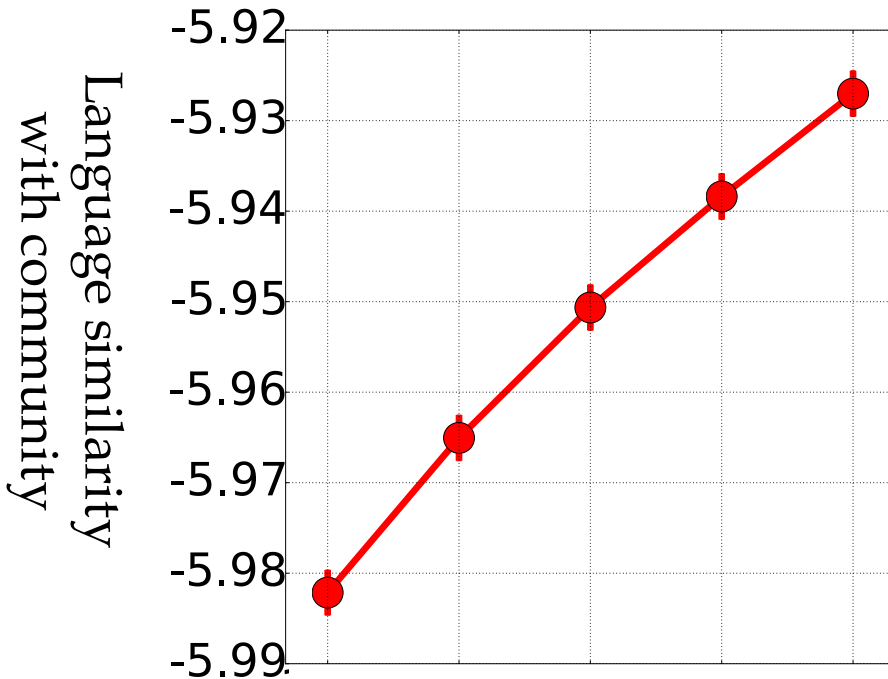
level of concentration, ⬇

visible community size, ⬇

community similarity ⬇

In our data, people do not settle down at all!

Users keep adopting each community's language

F: average language model score of $words_t$ vs $Community_t$

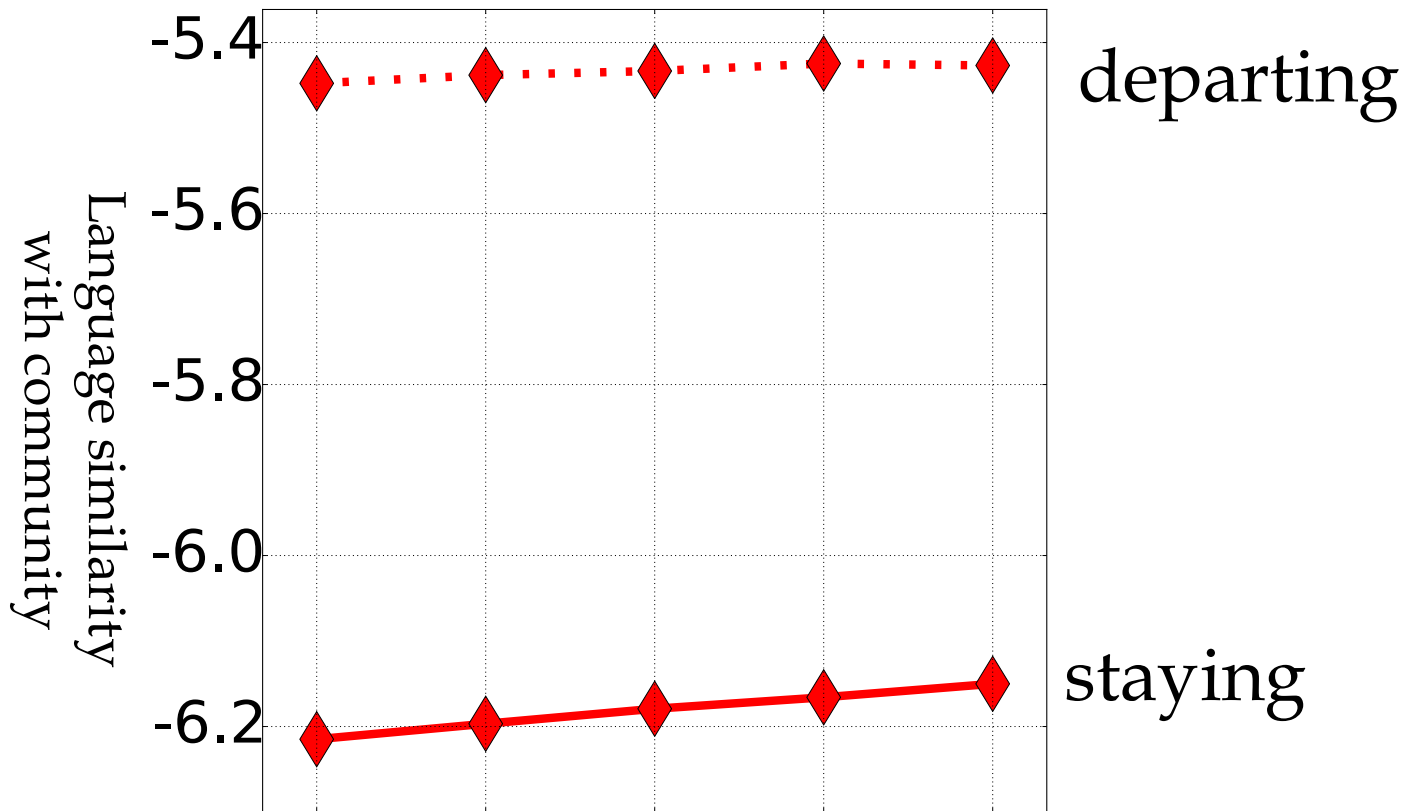A larger value indicates larger similarity

Users stay young:

Different from "users get old" (first increase and then drop) in single community setting

[Danescu-Niculescu-Mizil et al. 2013]

Error bars (tiny) show standard error.

Users keep adopting each community's language; future departing users **more** so

A larger value indicates larger similarity

Error bars (tiny) show standard error.

Using the first 50 posts to predict future departing status

- Feature sets
  - Baseline: average time gap [Danescu-Niculescu-Mizil et al. 2013, Dror et al. 2012, Yang et al. 2010]
  - Wandering pattern
  - Language
  - Evaluations
  - Combination of the above features
- Experiment details

  30 randomized train-test samples, logistic regression, F1 on departing users for evaluation

Features from trajectories outperform time-gap baseline

Users are "destined" to leave from the beginning!

Features from first 10 posts outperform baseline with all 50 posts

all

language

wandering

evaluations

baseline

Number of posts considered for prediction
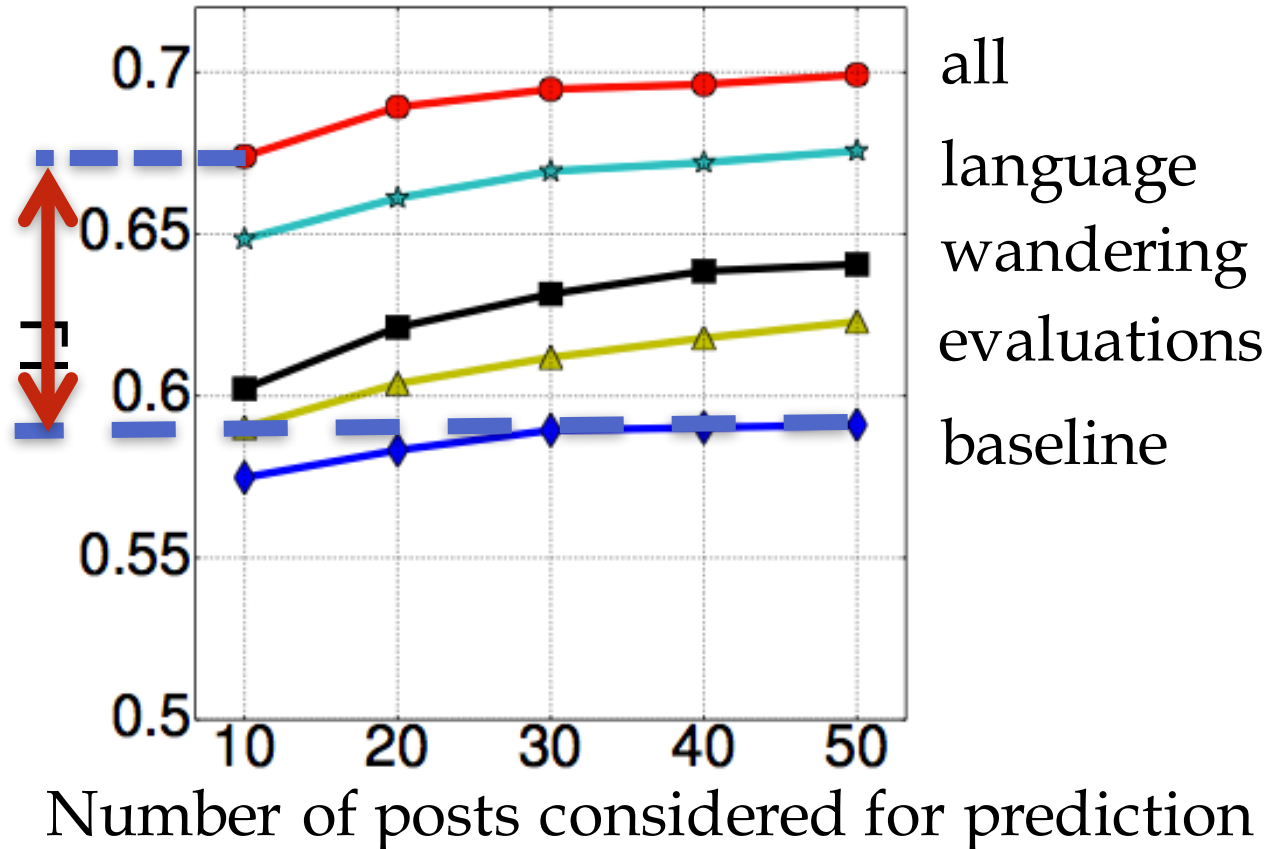
All differences along x-axis are significant (p<0.001) according to Wilcoxon signed rank test.

Macro level summary

- Design implications by investigating the multi-community setting:
  - First impressions matter
  - Give people choices to move to
- People, unlike trees, thrive by relocation（人挪活，树挪死）

Understand/predict human decisions
Improve social systems for humans

**Micro level:
effect of language**

**Macro level:
communities and networks**

cat

grumpycat

Effective speeches in
meetings and
group discussions

Community design: complementary
and substitutive effects

Effective explanation

TGIF! *TGIF!*

NO!

TGIF!

Community norm

Understand/predict human decisions
Improve social systems for humans

Micro level:
effect of language

Effective group discussions
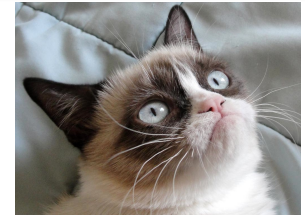Effective explanation

Macro level:
communities and networks

Community design: complementary & substitutive effects
Community norm

Meta level:
research methodology

- Experimentation + observational analysis
- New machine learning problems, e.g., human+machine, learning with justifications
- Theory testing

A journey to
a better understanding of humans
and a better world for humans!

Chenhao Tan
chenhao@chenhaot.com
https://chenhaot.com