USE OF SOCIAL NETWORKS IN RECOMMENDATION SYSTEMS

Tina Eliassi-Rad

Northeastern University

tina@eliassi.org

Roadmap

1. Measuring convergence and divergence of reading behaviors among friends
   - With Long T. Le [NewsKDD 2014]
2. A probabilistic model for using social networks in personalized item recommendation
   - With Allison Chaney and David Blei [RecSys 2015]

Measuring Convergence and Divergence of Reading Behaviors Among Friends

with Long T. Le

Appeared in NewsKDD 2014

Research Questions

- **Input:** Data from an online friendship network and its social reader*
- **[Q1]** How can we effectively capture the similarities between the reading behaviors of a user and her friends over time?
- **[Q2]** How can we effectively summarize such similarities across users?

---

* A reading application deployed on a social network

Motivation

- Better understand the activities on a social reader
- Use this newly gained understanding to devise better algorithms that promote application engagement

Challenges

- Heavy-tailed data
  - Some users / articles are extremely popular versus others are not

- Sparse data
  - Do not have enough observations for an article in a particular section

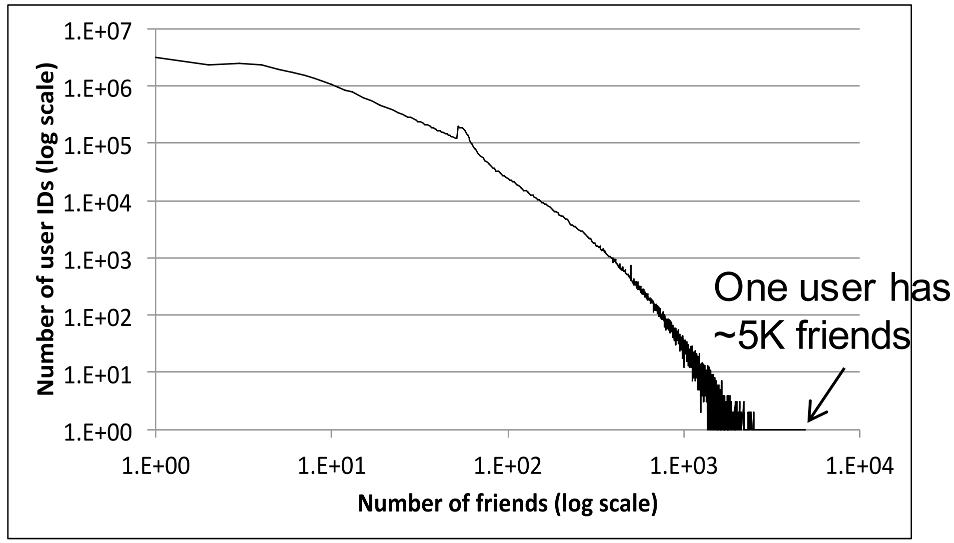
- A classic case of the "paradox of big data"

A Popular Social Reader

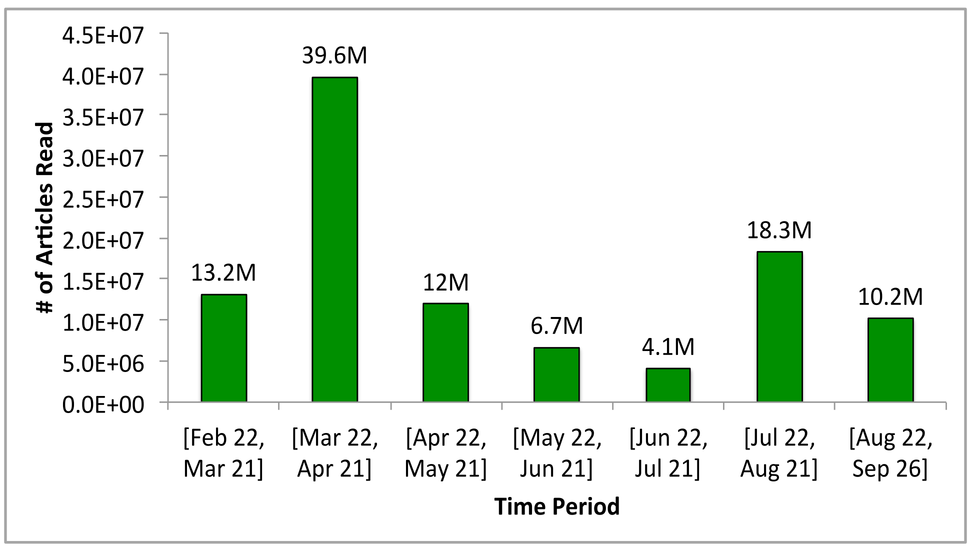
Our Data

- Friendship network (34GB)
  - 37.6M people
  - 502M friendship links
- Articles read (35.7GB)
  - 104M articles read from 2/2012 to 9/2012

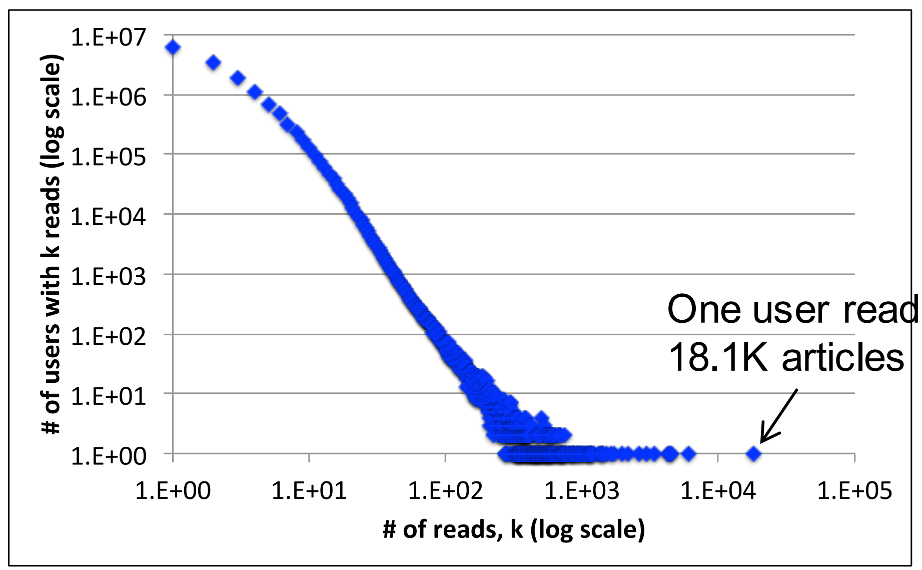
Degree Distribution of the Friendship Network



Distribution of Article Reads By Month

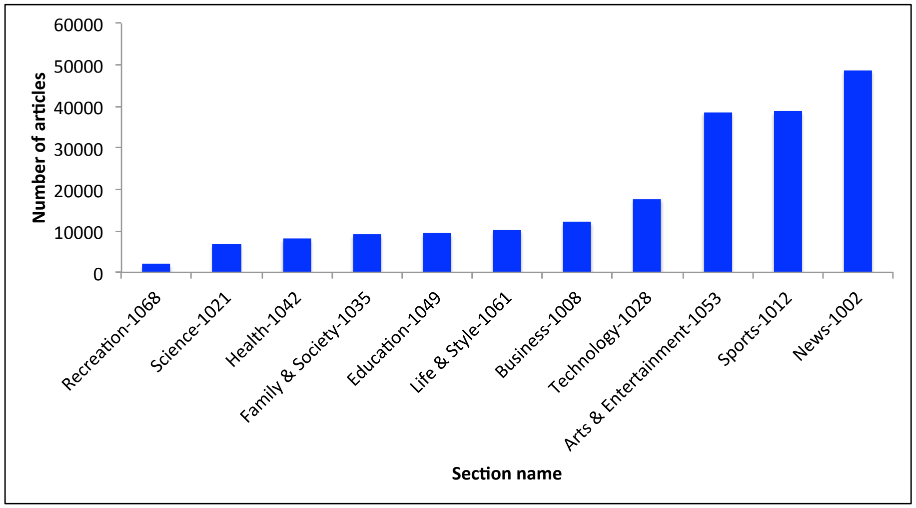


Article Reads Distribution

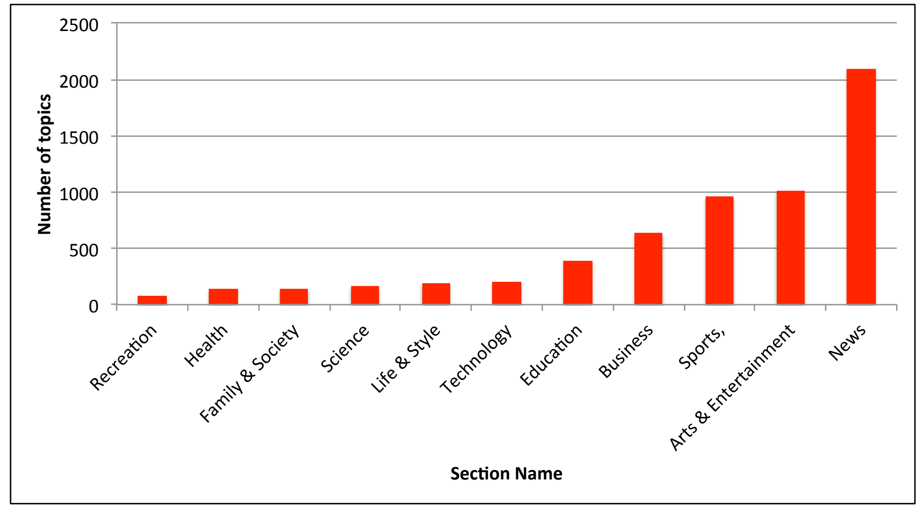


Articles, Topics, Sections

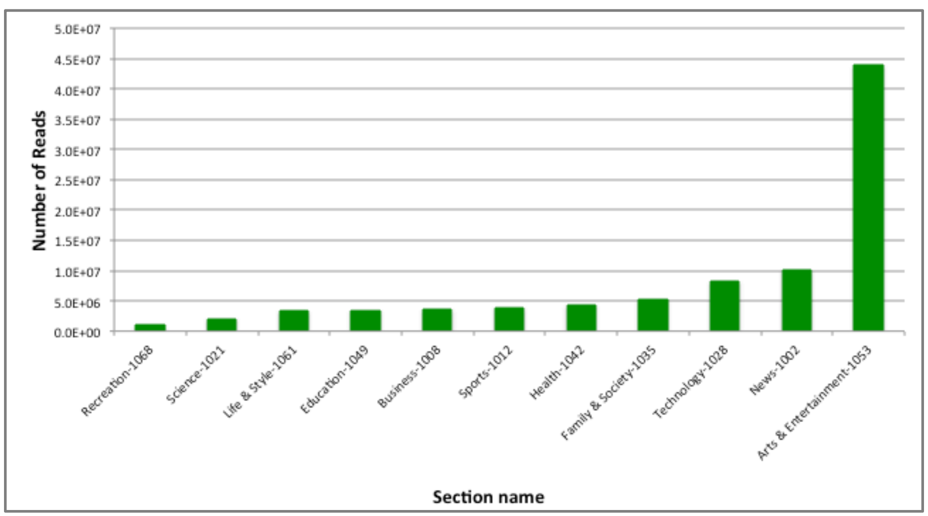
Distribution of Articles in Each Section



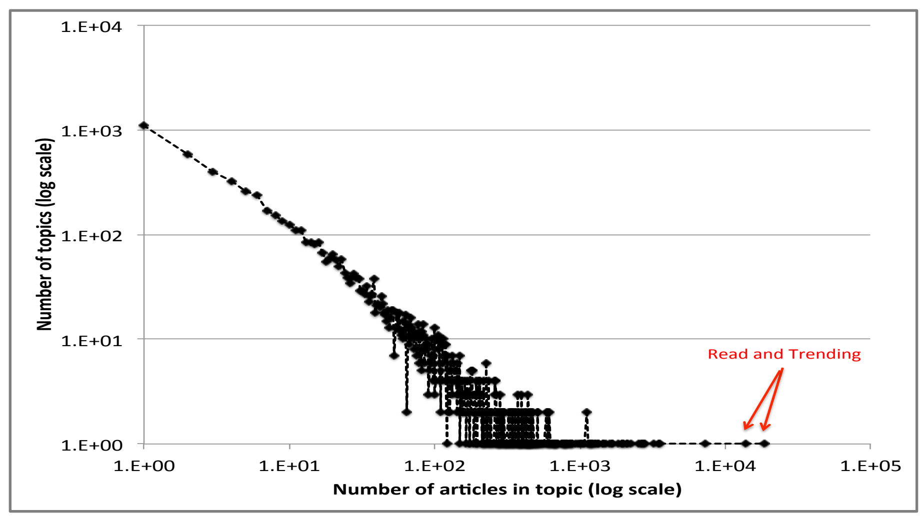
Distribution of Topics Across Sections



Distribution of Article Reads in Each Section

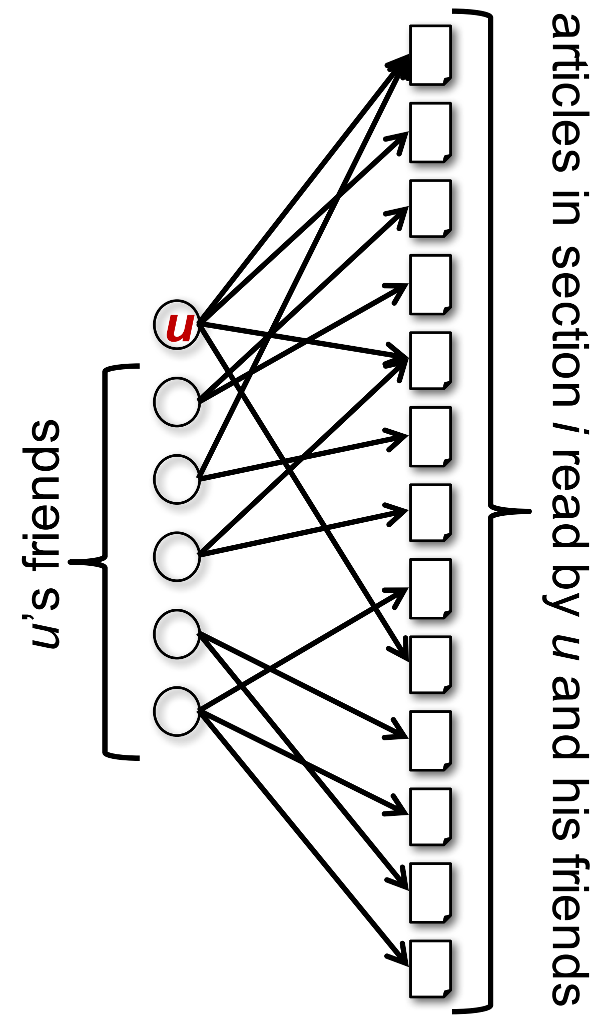


Distribution of Articles in Topics



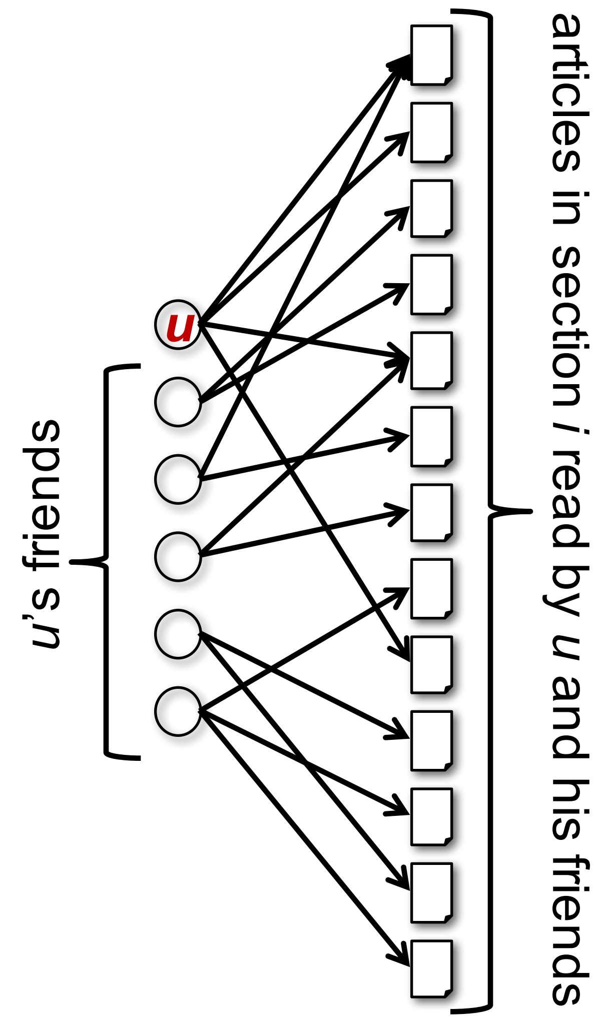
[Q1] How can we effectively capture the similarities between the reading behaviors of a user and her friends over time?

- Coverage
  - The amount by which the first-order Markov assumption holds between the reading behaviors of user *u* and her friends
- Divergence
  - The amount of inconsistency in the reading behaviors of user *u* and her friends across time



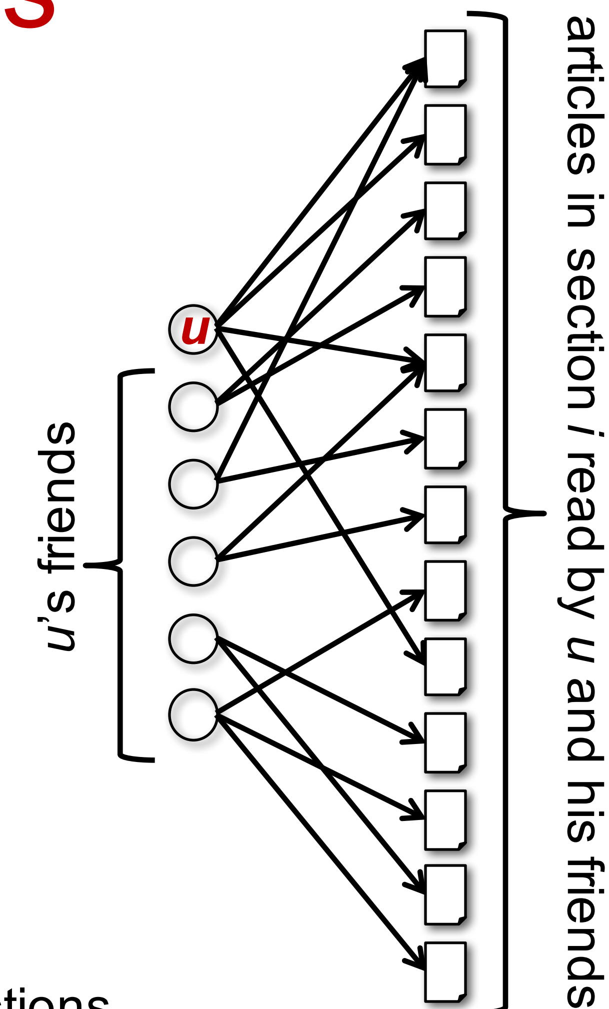
[Q2] How can we effectively summarize such similarities across users?

- Averaging across the users is not a good approach
  - Why?
  - Data is heavy-tailed and sparse
- Heavy-tailed
  - Some users / articles are extremely popular versus others are not
  - Articles in Arts & Entertainment are more popular than Science
- Sparsity
  - Do not have observations to compute coverage and divergence based on an article in a particular section



Graph Representations

- Eleven sections of the "newspaper"
  1. Recreation
  2. Health
  3. Family & Society
  4. Science
  5. Life & Style
  6. Technology
  7. Education
  8. Business
  9. Sports
  10. A&E
  11. News



- (User u + his friends) × Articles across all Sections
- (User u + his friends) × Articles in Section i
- (User u + his friends) × Topics across all Sections
- (User u + his friends) × Topics in Section i

Coverage & Divergence in Reading Behavior

- *Coverage* in reading tie-strength across time for *u* & friends

$$Cov(u, t_i, t_{i+1}) = \frac{\sum_{k=1}^{n_u} t_i[u,k] \times TS(u,k,t_{i+1})}{\sum_{k=1}^{n_u} TS(u,k,t_{i+1})}$$

where $t_i[u, k] = 1$ when $TS(u, k, t_i) > 0$

- *Divergence* in reading tie-strength across time for *u* & friends

$$Div(u, t_i, t_{i+1}) = \frac{1}{n_u} \sum_{k=1}^{n_u} \frac{|TS(u,k,t_i) - TS(u,k,t_{i+1})|}{|TS(u,k,t_i)| + |TS(u,k,t_{i+1})|}$$

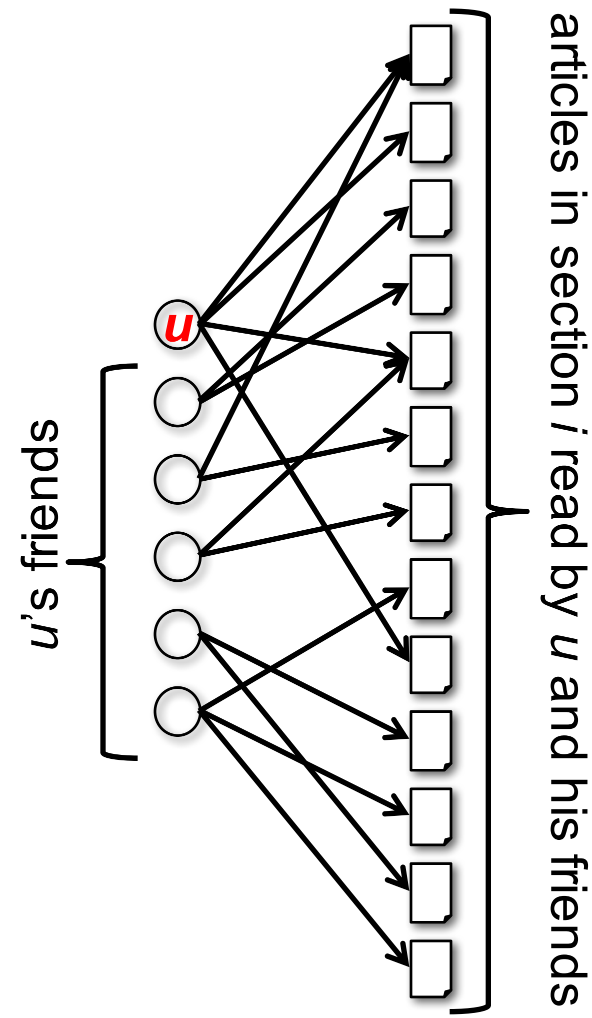
n_u = # of u's friends

Summarizing Coverage & Divergence Values Across All users

- Dealing with heavy-tailed data
  - More popular users should get more weight
  - More popular articles should get more weight
- Solution
  - Compute coverage and divergence on a *big table*
    - Rows = friendship pairs for a given section (i.e., a triple ⟨*u*, friend of *u*, section *s*⟩)
    - Columns = tie strengths of each triple at different time periods

Summarizing Coverage & Divergence Values Across All users

- Dealing with sparsity in data
  - Utilize bipartite graphs whose article nodes are articles-read across all sections
  - Computation at a coarser level
    - Will produce a higher coverage and lower divergence values
    - **But** they allow us to compute coverage and divergence for everyone



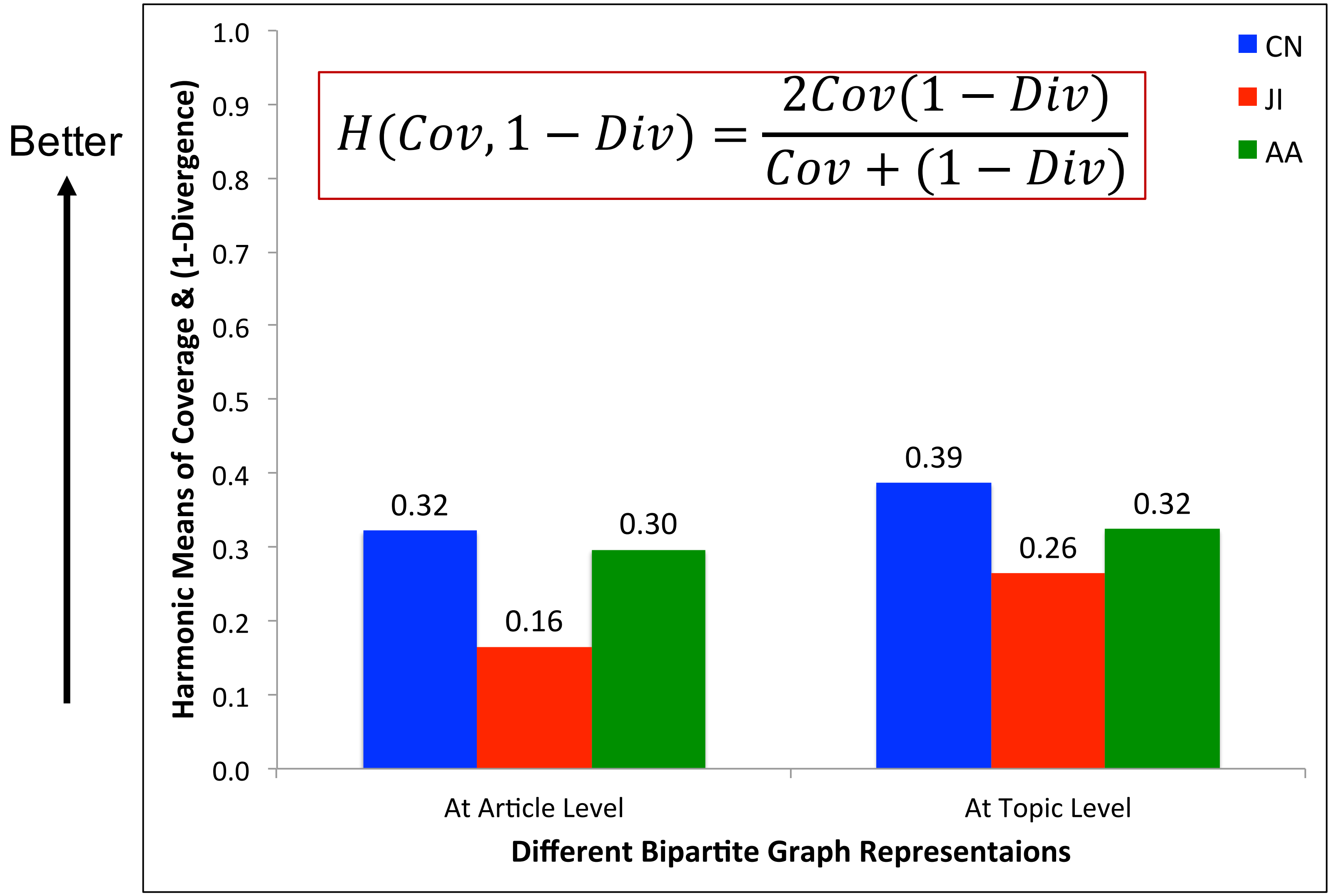
Tie-strength Measures

- Common neighbor (CN)
  - # of common articles that both *u* and *v* read
- Jaccard Index (JI)
  - Similar to CN
  - Normalizes for how “social” *u* and *v* are
- Adamic-Adar (AA):
  - Tie strength increases as # of common (read) articles increases
  - Tie strength for a common (read) article is 1 over log of the # of individuals who read that article

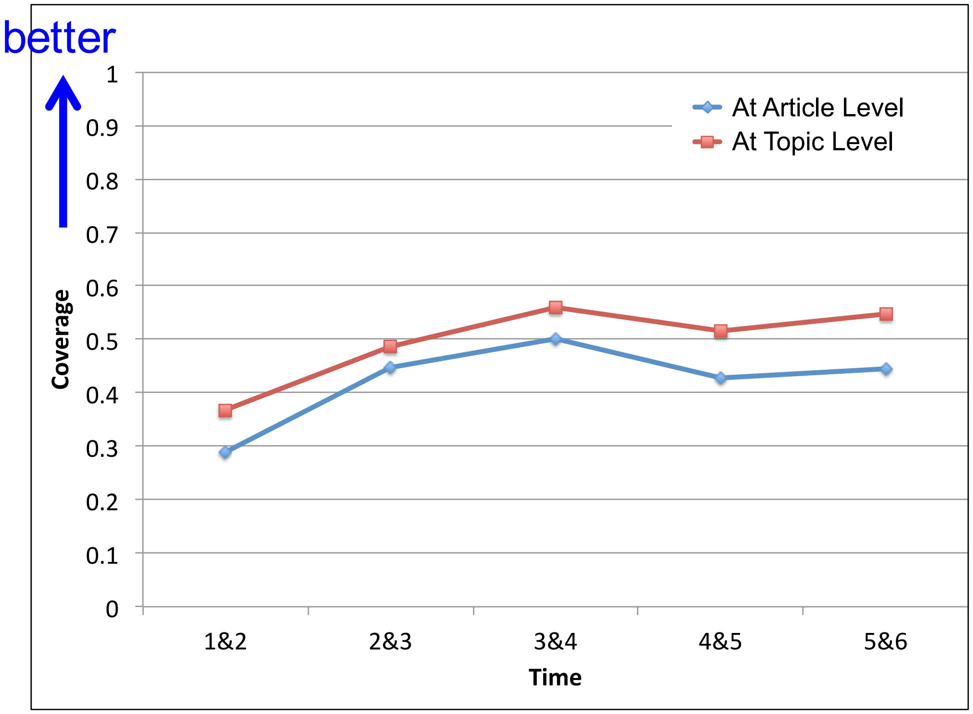
---

Mangesh Gupte, Tina Eliassi-Rad: Measuring tie strength in implicit social networks. *WebSci* 2012: 109-118

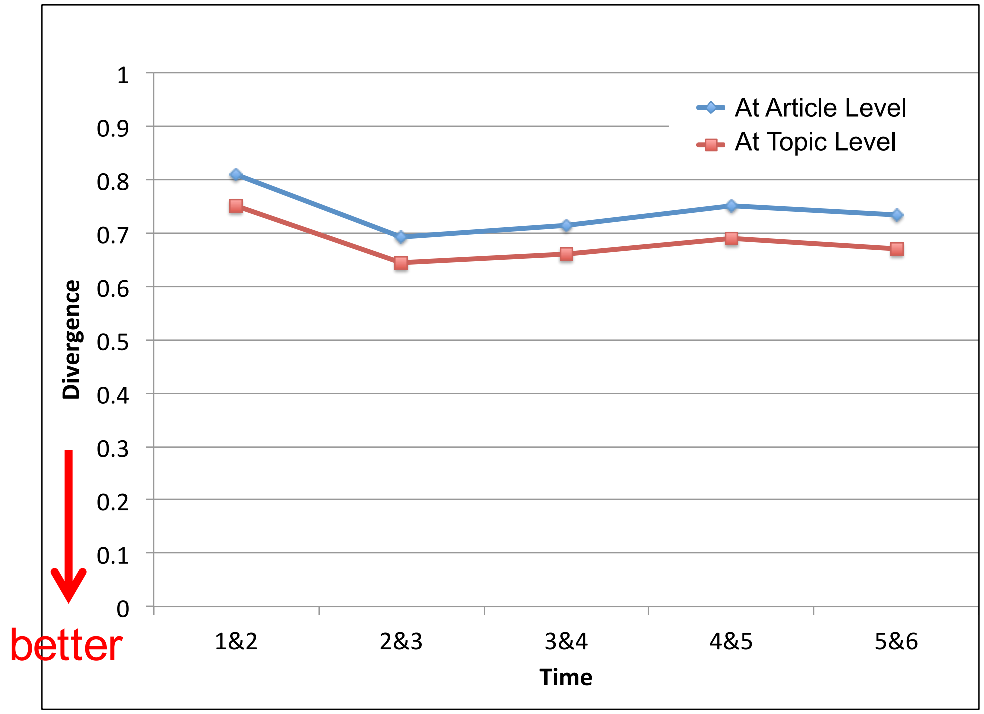
CN is Better Than JI and AA in capturing Coverage and Divergence



Topic Level Better Than Article Level



Coverage
Of **Common Neighbor**
Over Time

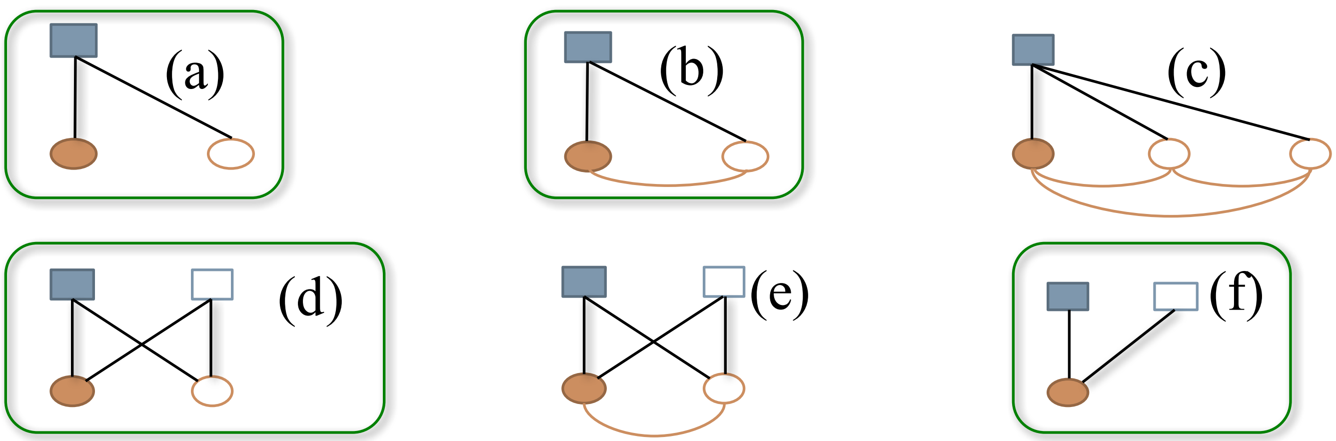


Divergence
of **Common Neighbor**
Over Time

Testing Socialness of Social Reader

- Test multiple social theories via Exponential Random Graph Models (ERGMs) in a longitudinal study and in a cross-sectional study

$$p(\mathbf{Y} = \mathbf{y}|\boldsymbol{\theta}) = \frac{1}{Z} e^{\boldsymbol{\theta}^{\top} \boldsymbol{\phi}(\mathbf{y})}$$



(a) Preferential Attachment; (b) & (c) Social Influence and Contagion Theories; (d) & (e) Cognitive Consistency and Balance Theories; (f) Individual Reading Tendency

With Mengxiao Zhu (ETS Research)

Social Reader Analysis: Summary

- A case-study on reading activities on the WaPo Social Reader
- How can we effectively capture the similarities between the reading behaviors of a user and her friends over time?
  - Coverage: amount by which the first-order Markov assumption holds between reading behaviors of user *u* and her friends
  - Divergence: amount of inconsistency in their reading tie-strength across time
- How can we effectively summarize such similarities across users?
  - Compute coverage & divergence on ⟨u, friend of u, section s⟩ triples across time; and use the taxonomy over articles
- Take-away points from the experiments:
  - Common neighbor is better than Jaccard Index and Adamic-Adar in this application (concurs with Gupte & Eliassi-Rad [WebSci'12])
  - Operate at the coarser topic level

A Probabilistic Model for Using Social Networks in Personalized Item Recommendation

with Allison Chaney and David Blei

Appeared in RecSys 2015

Personalized Item Recommendation

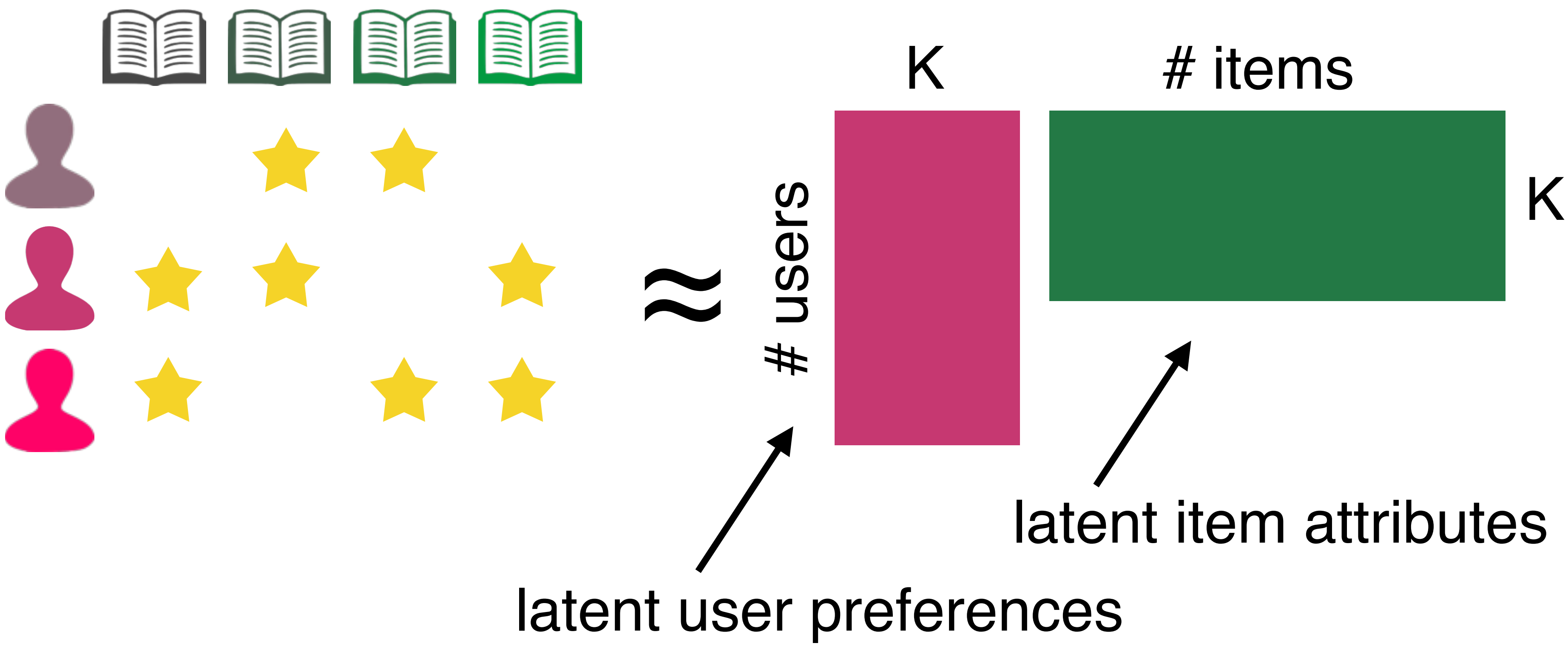
Anna Karenina

Winter’s Tale

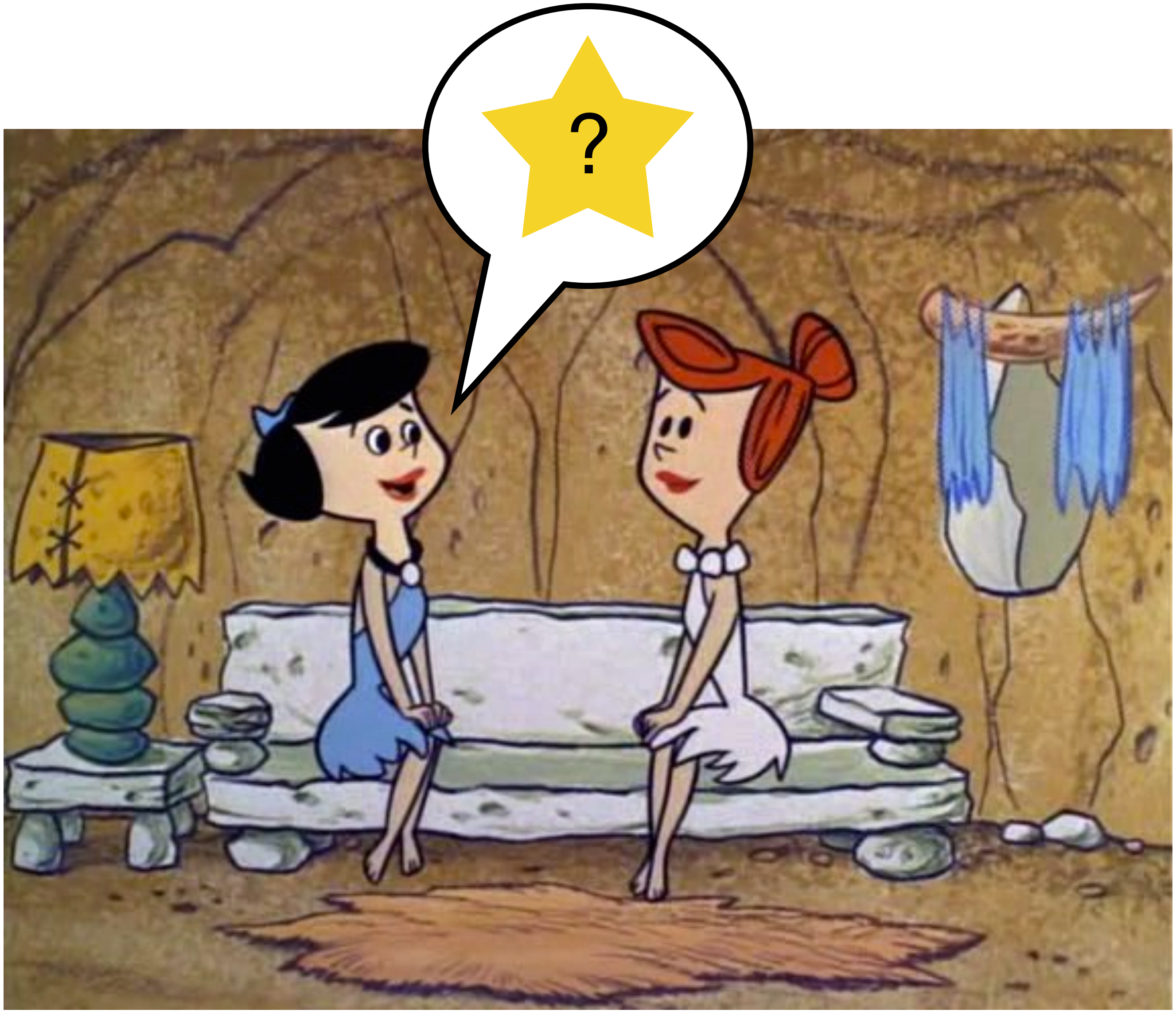
East of Eden

???

Matrix Factorization



Including Social Networks

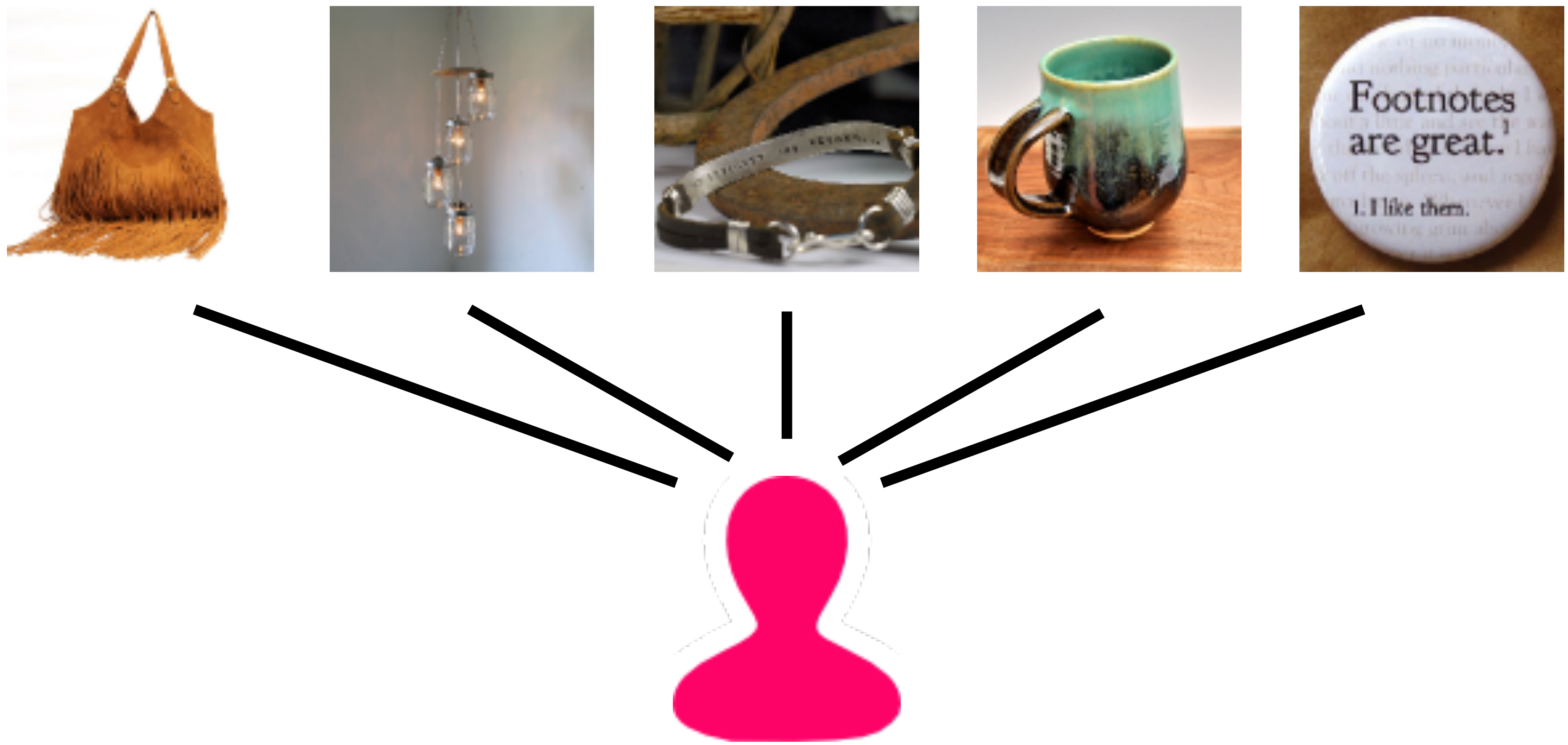


Including Social Networks

- Matches our intuition
- Introduces explainable serendipity
- Improves performance
- Helps us learn about the social network

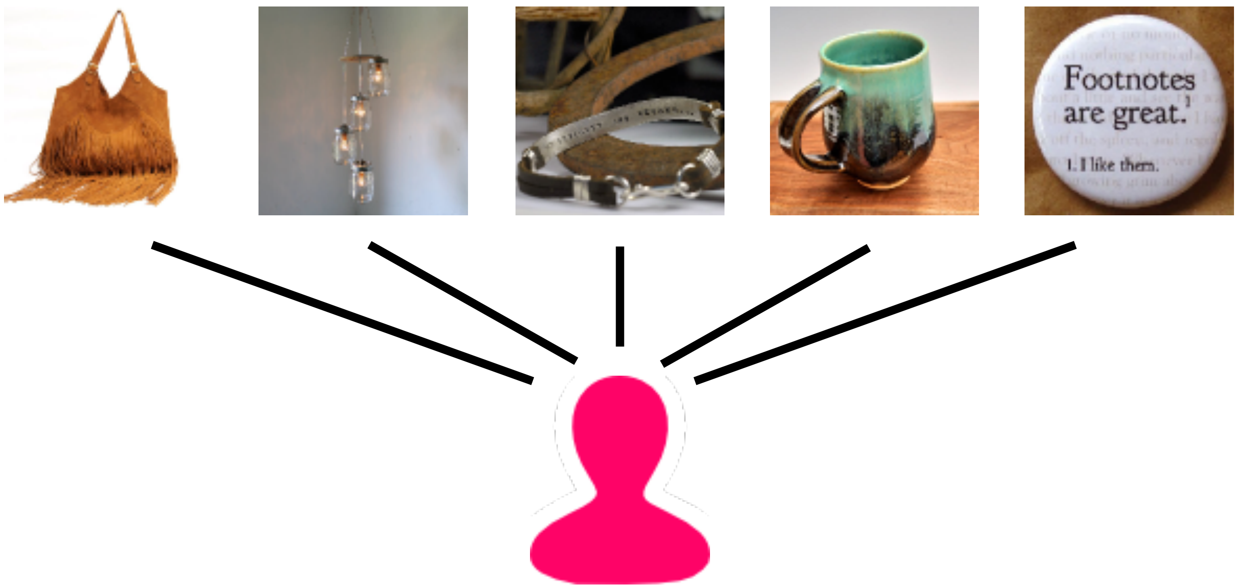
An Example Etsy User

An Example Etsy User



Edge = user *u* clicked on item *i*

An Example Etsy User



An Example Etsy User

An Example Etsy User

An Example Etsy User

An Example Etsy User

Overview

observed data

ratings

network

inference algorithm

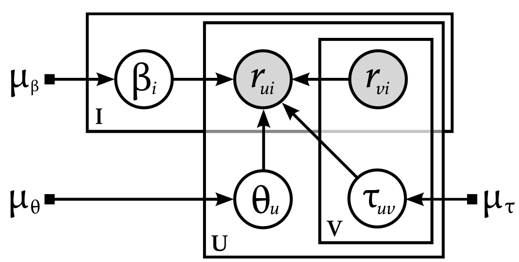
learned parameters

item attributes

user preferences

user influence

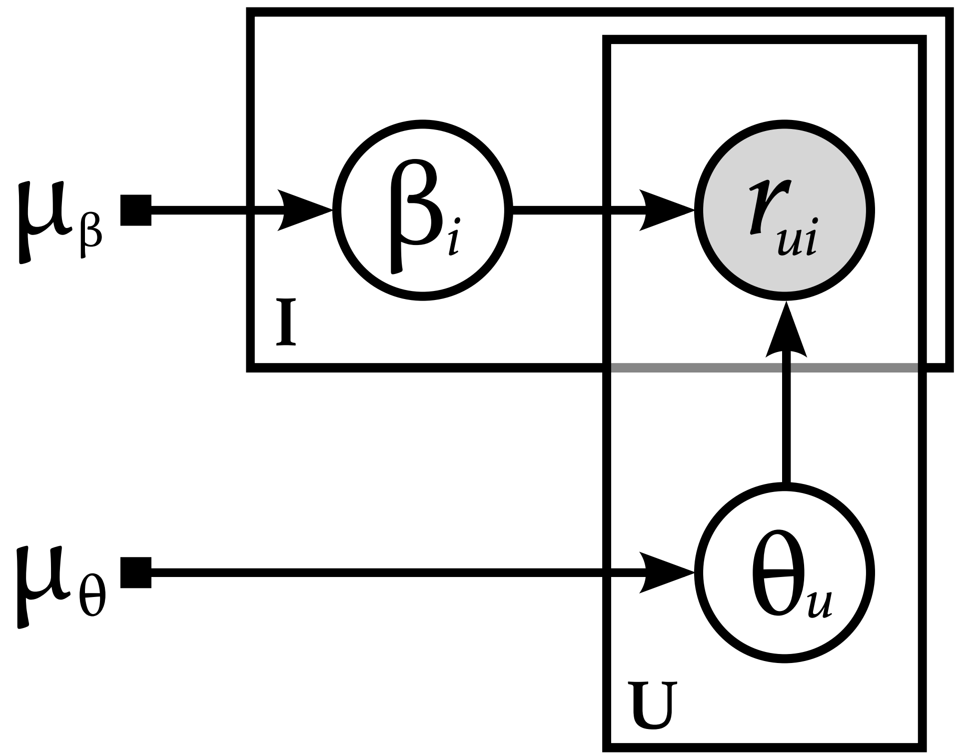
model assumptions



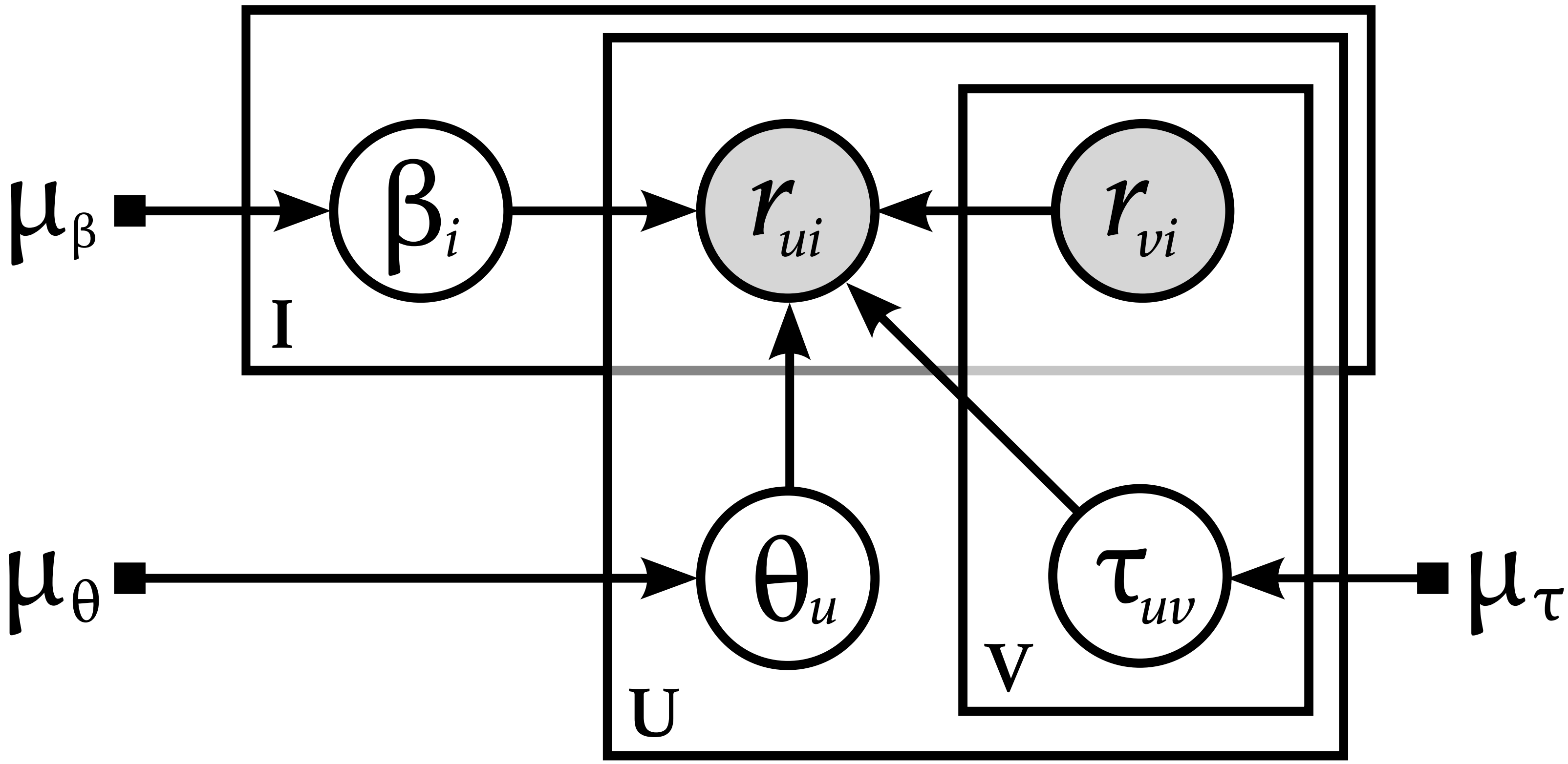
recommendations

Matrix Factorization

$$r_{ui} \sim \mathrm{Poisson}(\theta_u^{\mathrm{T}} \beta_i)$$

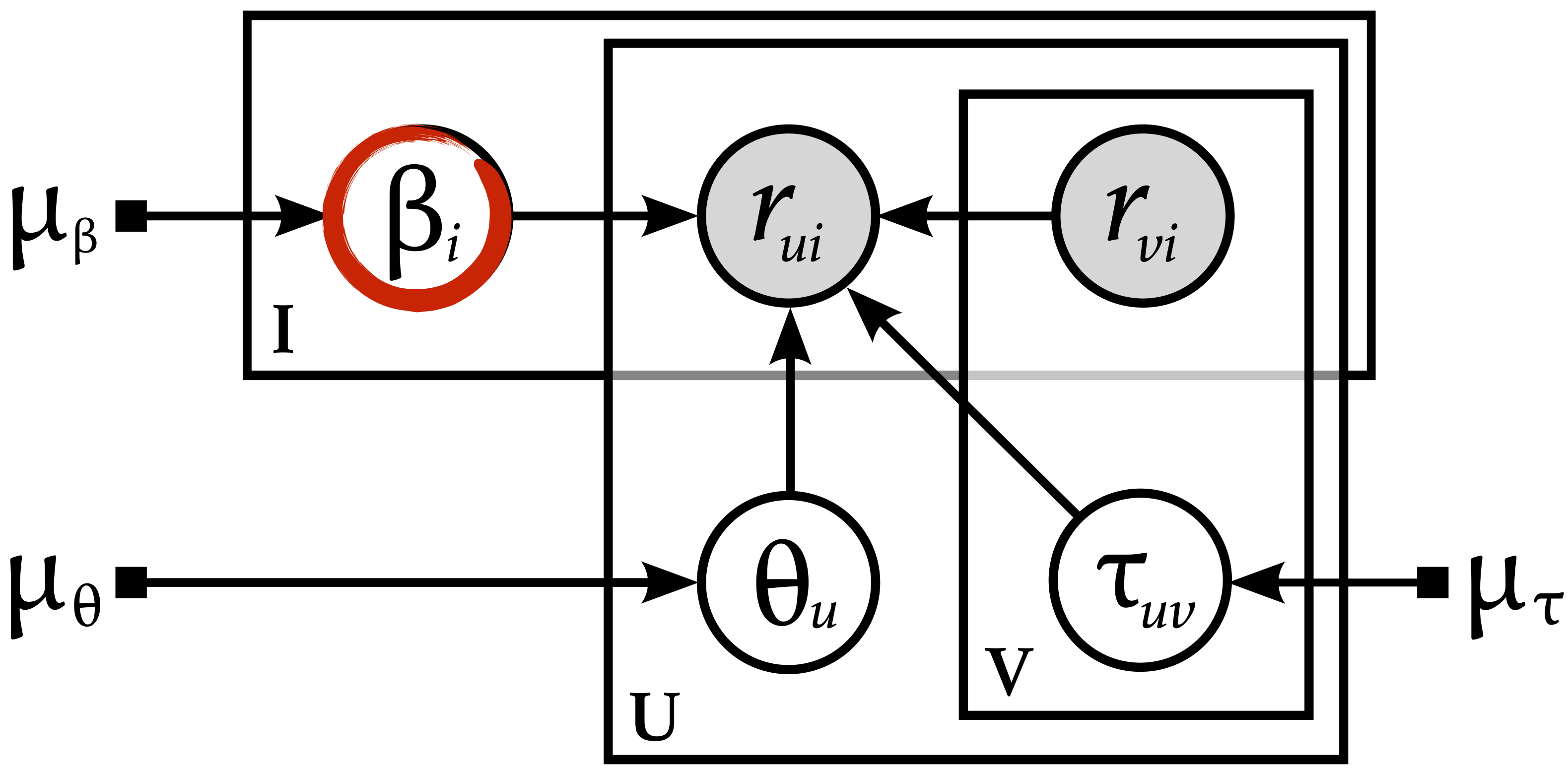


Social Poisson Factorization



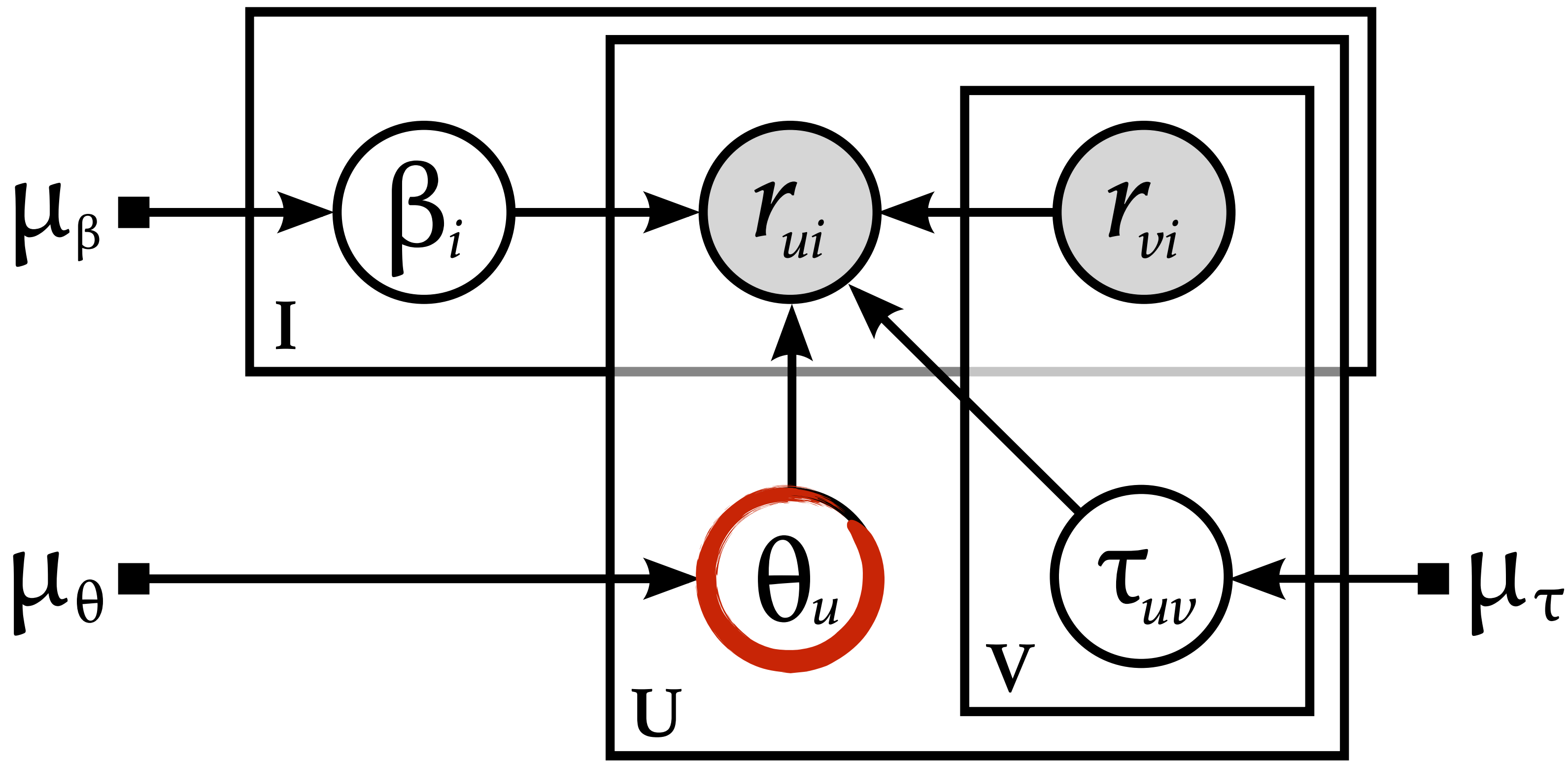
Item Attributes

$$\beta_{ik} \sim \text{Gamma}(a_\beta, b_\beta)$$



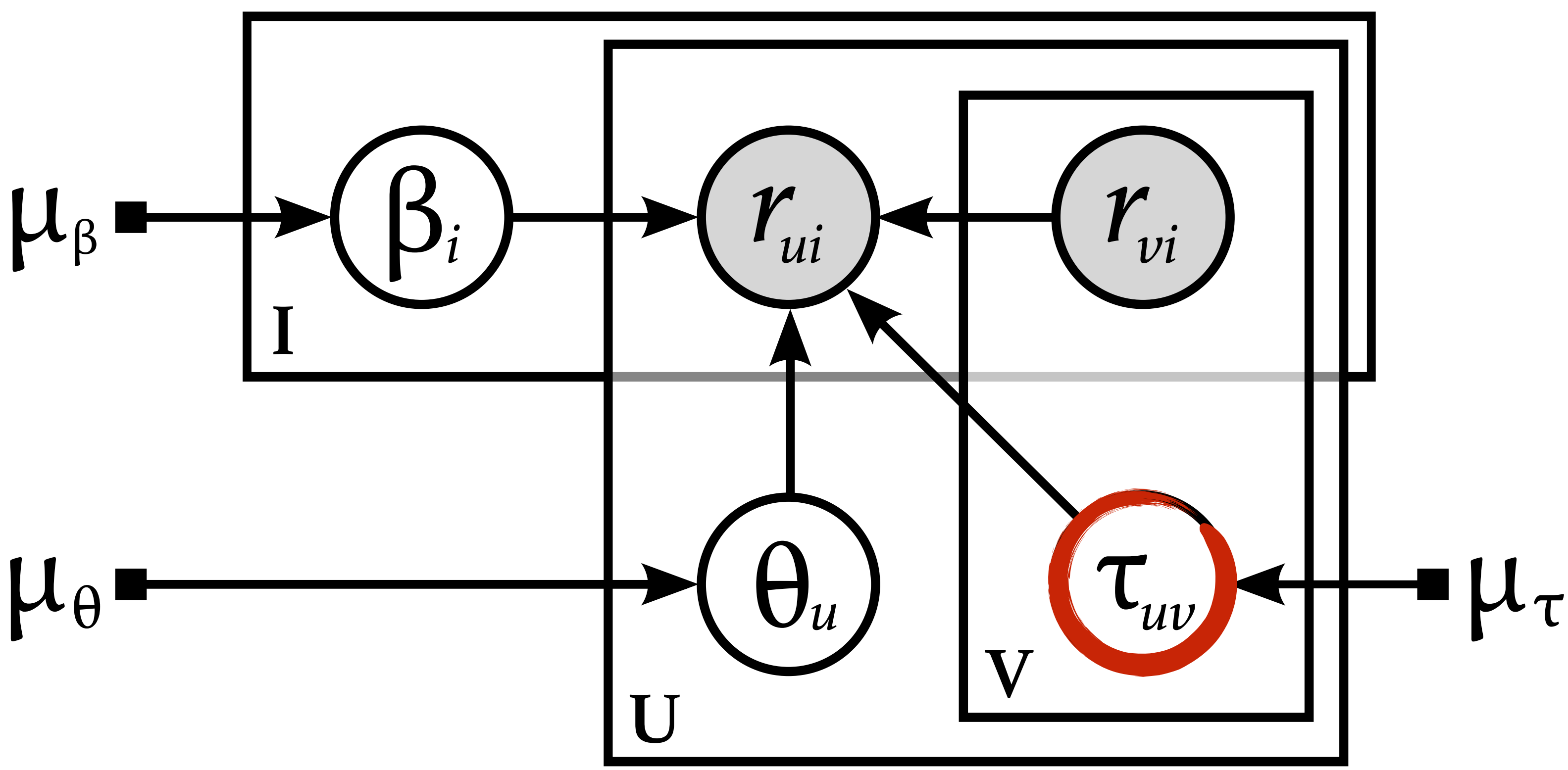
User Preferences

$$\theta_{uk} \sim \mathrm{Gamma}(a_\theta, b_\theta)$$

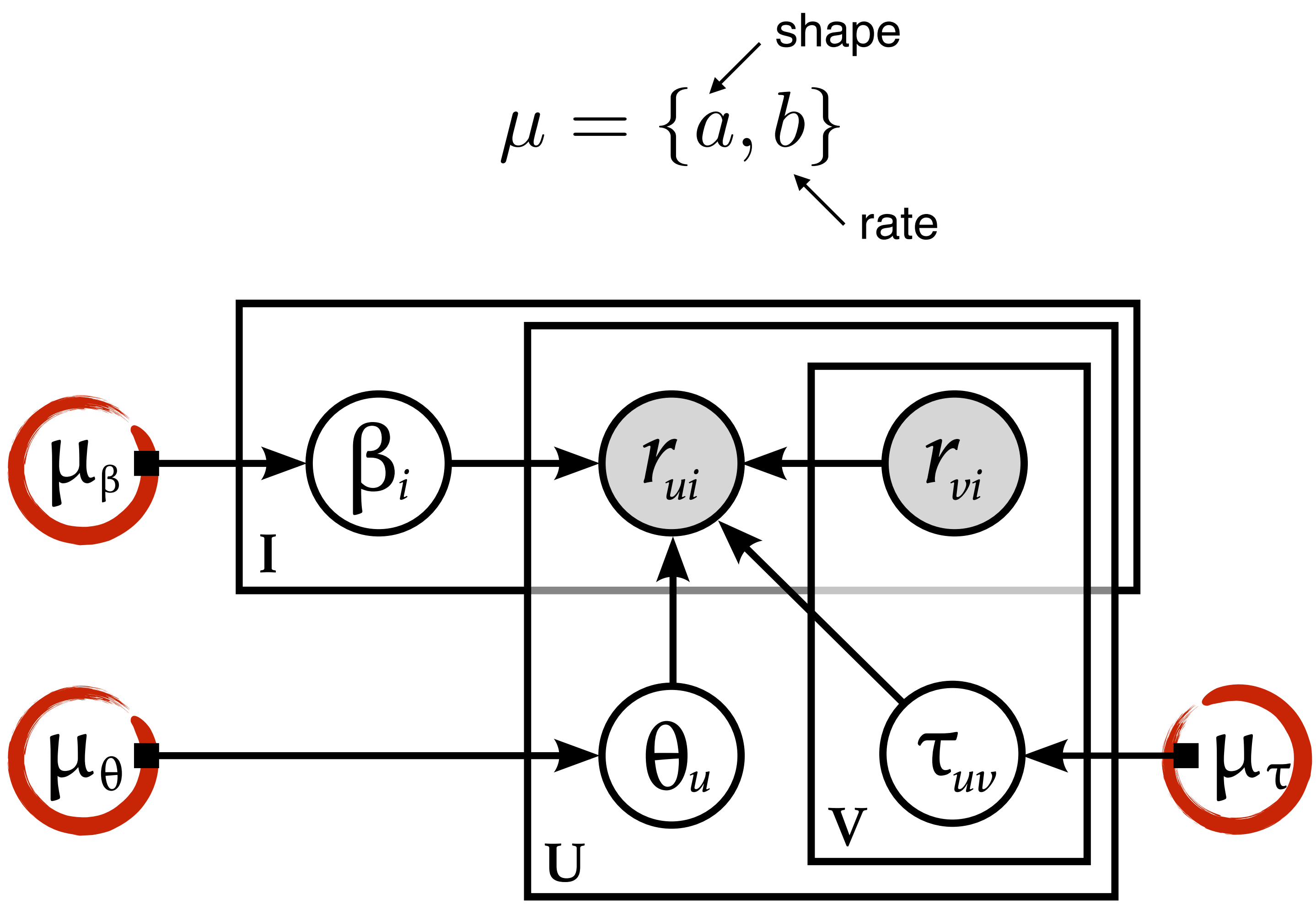


User Influence

$$\tau_{uv} \sim \mathrm{Gamma}(a_\tau, b_\tau)$$

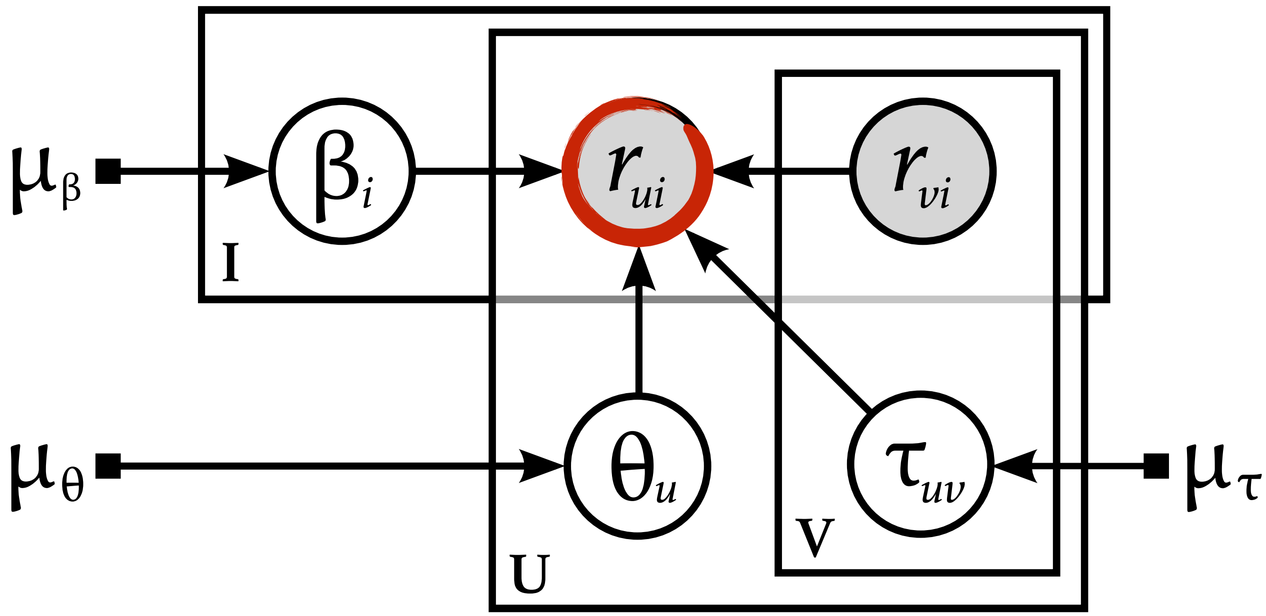


Hyperparameters



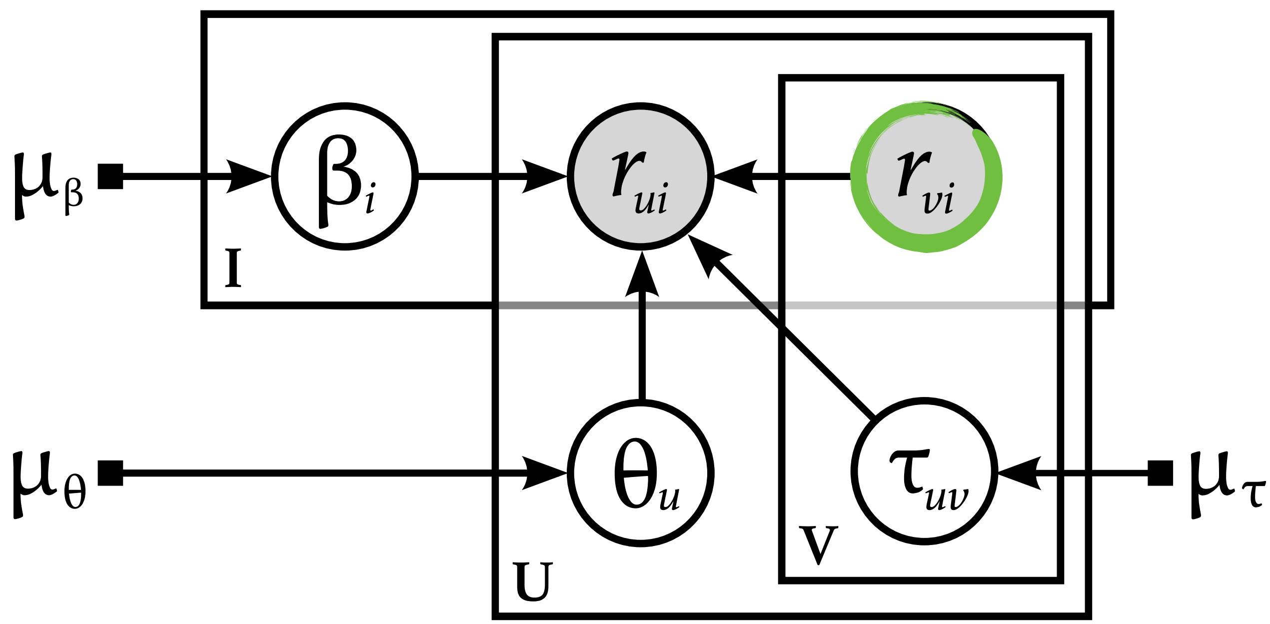
Ratings

$$r_{ui} \mid r_{-u,i} \sim \text{Poisson}\left(\theta_u^\top \beta_i + \sum_{v \in N(u)} \tau_{uv} r_{vi}\right)$$



Ratings

$$r_{ui} \mid r_{-u,i} \sim \text{Poisson}\left(\theta_u^\top \beta_i + \sum_{v \in N(u)} \tau_{uv} r_{vi}\right)$$



Posterior Inference

How do we go from a generative model to finding the values of the variables that best fit our data?

Posterior Distribution

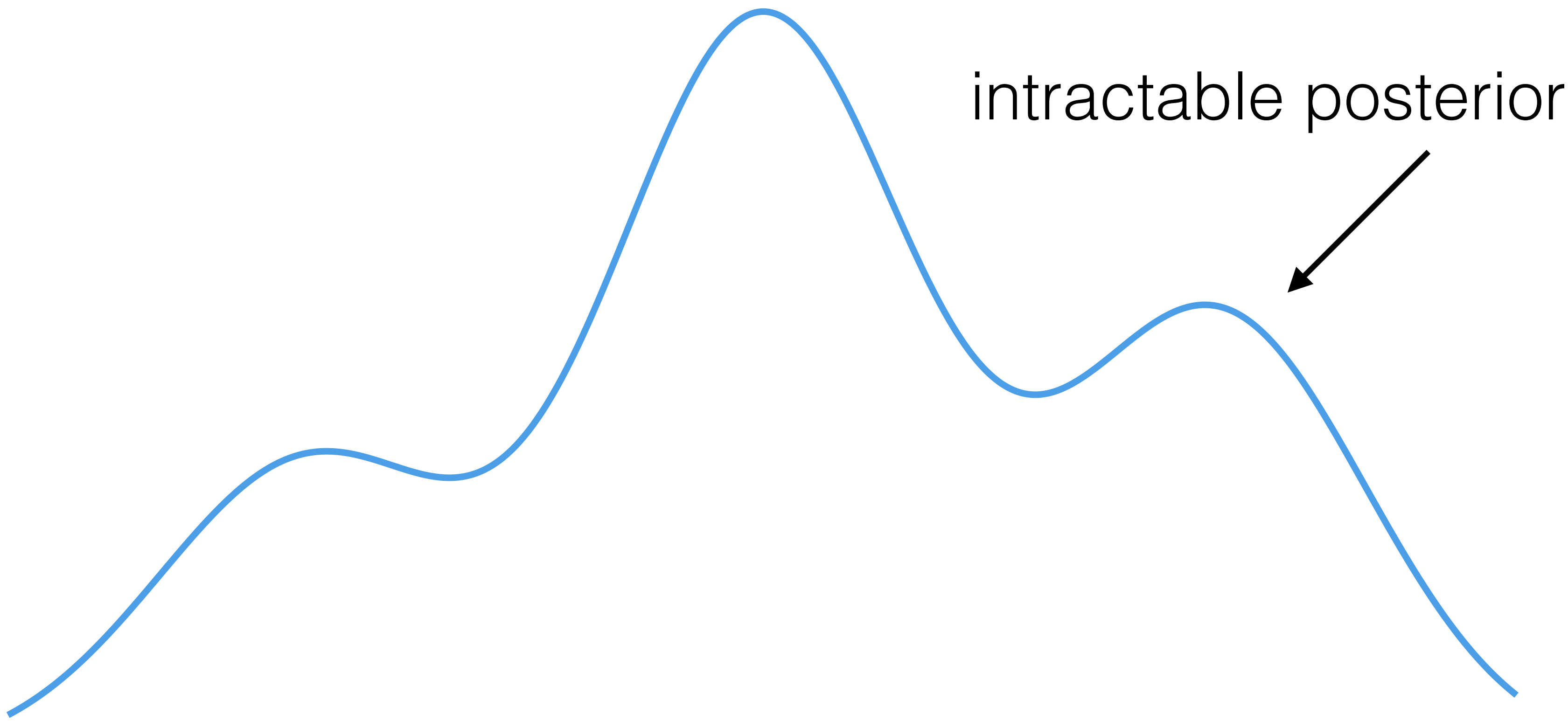
latent model parameters

$$p(\beta, \theta, \tau \mid \mathbf{R}, \mathbf{N}, \mu) = \frac{p(\beta, \theta, \tau, \mathbf{R}, \mathbf{N} \mid \mu)}{\int_\beta \int_\theta \int_\tau p(\beta, \theta, \tau, \mathbf{R}, \mathbf{N} \mid \mu)}$$

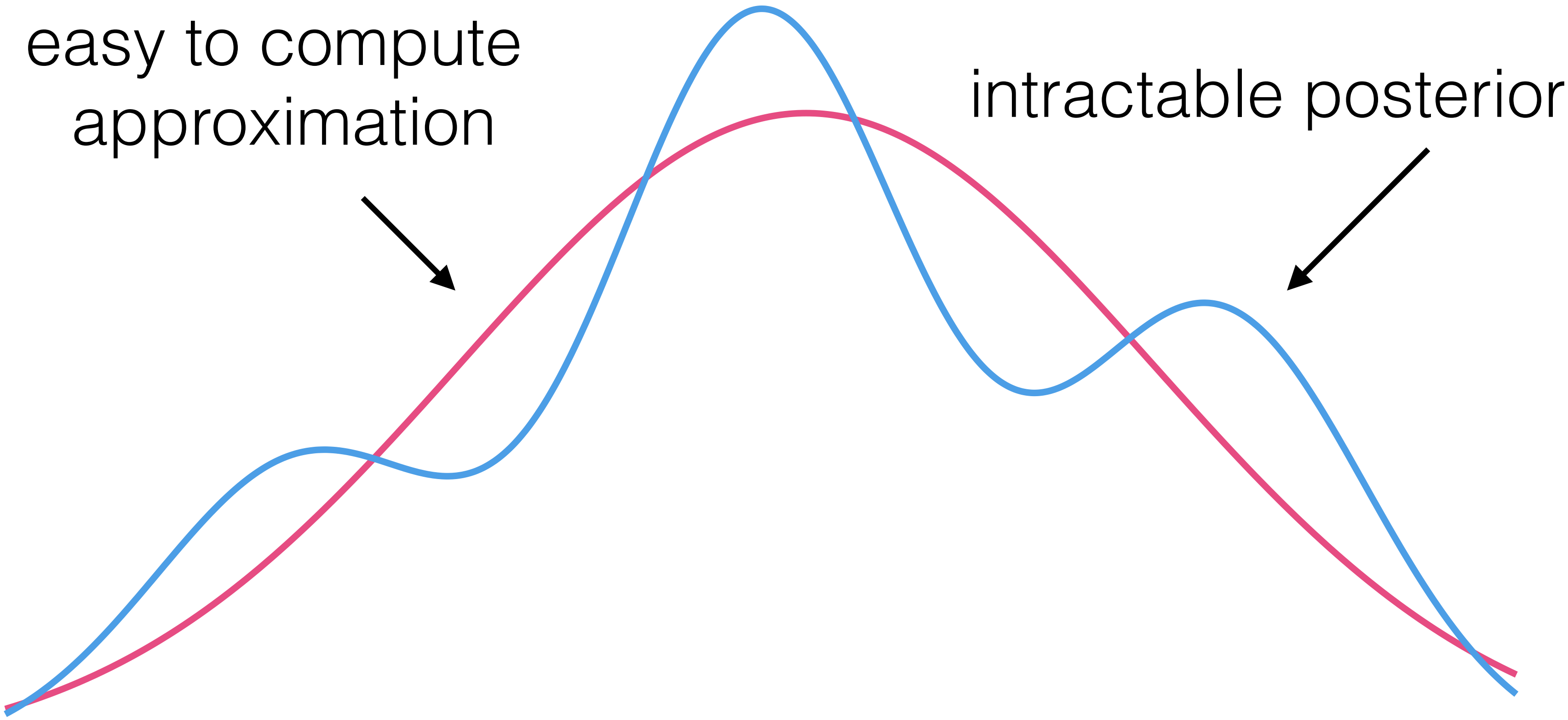
observed data

model hyperparameters

Mean Field Variational Inference



Mean Field Variational Inference



Recommendation

$$\mathbf{E}[r_{ui}] = \mathbf{E}[\theta_u]^\top \mathbf{E}[\beta_i] + \sum_{v \in N(u)} \mathbf{E}[\tau_{uv}] r_{vi}$$

Data

source	# users	# items	% ratings	% edges
Ciao	7,000	98,000	0.038%	0.103%
Epinions	39,000	131,000	0.012%	0.011%
Flixster	132,000	42,000	0.122%	0.006%
Douban	129,000	57,000	0.221%	0.016%
Social Reader	122,000	6,000	0.065%	0.001%
Etsy	40,000	5,202,000	0.009%	0.300%

etsy.com and librec.net/datasets.html

Comparison Approaches

SoRec	Ma et al., SoRec: Social Recommendation Using Probabilistic Matrix Factorization, *SIGIR* 2008.
RSTE	Ma et al., Learning to Recommend with Social Trust Ensemble, *SIGIR* 2009.
SocialMF	Jamali and Ester, A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks, *RecSys* 2010.
TrustMF	Yang et al., Social Collaborative Filtering by Trust, *IJCAI* 2013.
TrustSVD	Guo et al., TrustSVD: Collaborative Filtering with Both the Explicit and Implicit Influence of User Trust and of Item Ratings, *AAAI* 2015.

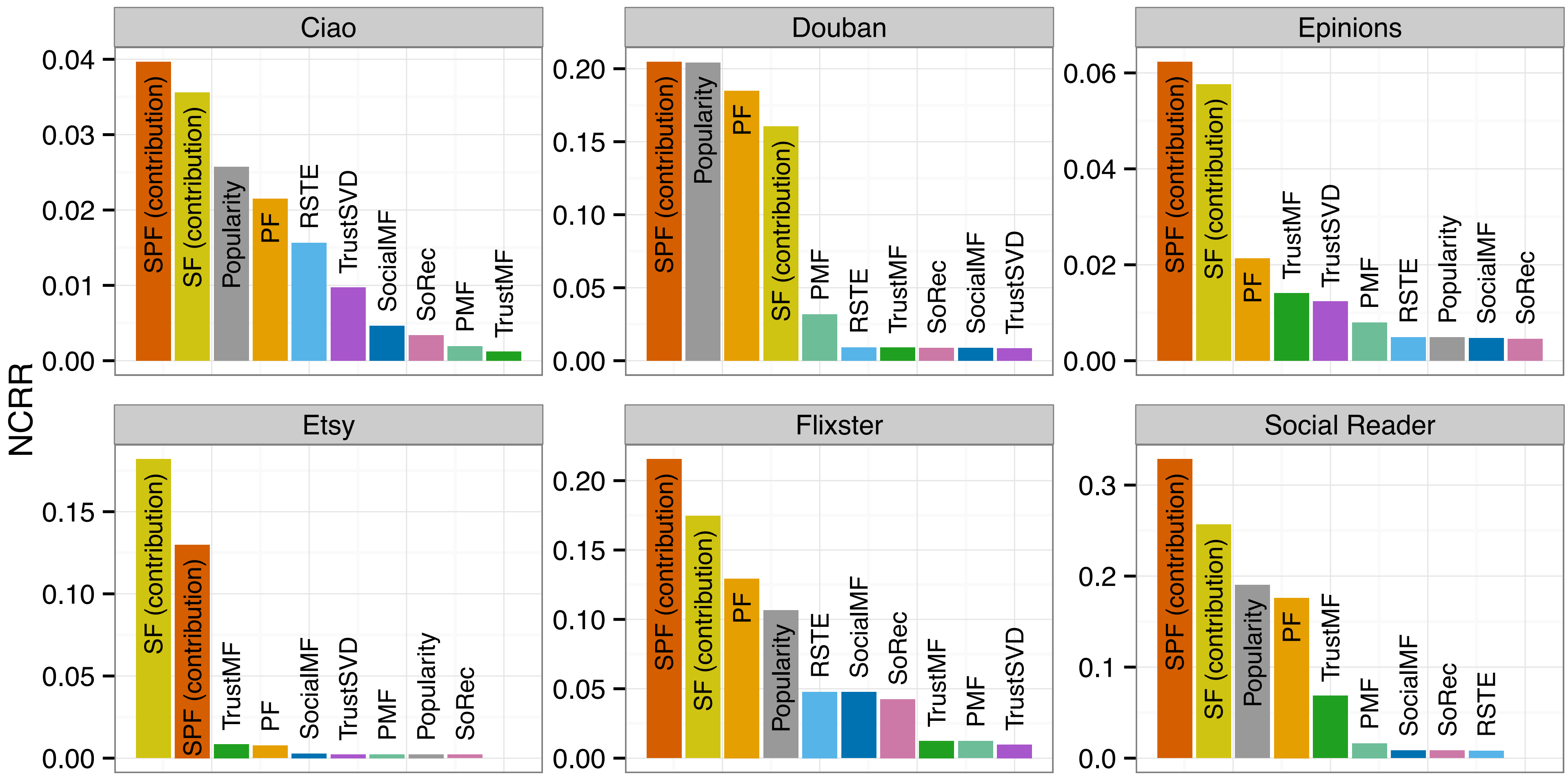
librec.net

Evaluation on Held-out Data

$$CRR(user) = \sum_{n=1}^{N} \frac{\mathbf{1}[rec_n \in \mathcal{H}]}{n} = \sum_{i \in \mathcal{H}} \frac{1}{rank(i)}$$

$$NCRR(user) = \frac{CRR(user)}{\text{ideal } CRR(user)}$$

Results



NCRR
Ciao
SPF (contribution)
SF (contribution)
Popularity
PF
RSTE
TrustSVD
SocialMF
SoRec
PMF
TrustMF
Douban
SPF (contribution)
Popularity
PF
SF (contribution)
PMF
RSTE
TrustMF
SoRec
SocialMF
TrustSVD
Epinions
SPF (contribution)
SF (contribution)
PF
TrustMF
TrustSVD
PMF
RSTE
Popularity
SocialMF
SoRec
Etsy
SF (contribution)
SPF (contribution)
TrustMF
PF
SocialMF
TrustSVD
PMF
Popularity
SoRec
Flixster
SPF (contribution)
SF (contribution)
PF
Popularity
RSTE
SocialMF
SoRec
TrustMF
PMF
TrustSVD
Social Reader
SPF (contribution)
SF (contribution)
Popularity
PF
TrustMF
PMF
SocialMF
SoRec
RSTE

Social Poisson Factorization: Summary

- Performs better than comparison models
- Is interpretable and has explainable serendipity
- Scales well to large data
- Source code available at ajbc.io/spf

Open Issues to be Addressed

- Topical influence
- Include timestamps on users’ behavior
- A/B testing
- Explore biases in data and how to correct for them

Roadmap

1. Measuring convergence and divergence of reading behaviors among friends
   - With Long T. Le [NewsKDD 2014]
2. A probabilistic model for using social networks in personalized item recommendation
   - With Allison Chaney and David Blei [RecSys 2015]

Thank You

- Papers at http://eliassi.org/pubs.html
- Contact me at tina@eliassi.org
- Supported by NSF, DTRA, DARPA, LLNL, and WaPo Labs