

Introduction to Social Computing

Social Recommendation

Irwin King

ATT Labs, Research
&

Department of Computer Science and Engineering
The Chinese University of Hong Kong

king@cse.cuhk.edu.hk
<http://www.cse.cuhk.edu.hk/~king>

©2010 Irwin King. All rights reserved.



Outline

- **Social Search Engine**
- Social Recommender Systems
- Social Media Analysis



Macro Definition

- Search in
 - Shared bookmarks
 - Collaborative directories
 - Collaborative news/opinions
 - Social Q&A sites
 - etc...



Micro Definition

Leveraging Your Social Networks for Searching



Leveraging All Kinds of Web Accounts



Google's Social Search

[Results from people in your social circle for google bus](#)

[Google Maps Ad on Chicago Bus - Googlified](#)



[haochi](#) - connected via Tom on digg.com

google transit chicago bus ad. Google Transit recently became available to Chicago users and the Chicago team has been very active in ...

googlified.com/google-maps-ad-on-chicago-bus/

[More results from haochi »](#)

[Google Student Blog: The Google Apps Bus stops at the beginning](#)



Google Students - connected via twitter.com

Almost two years later, the **Google App to School bus** pulled into Arizona State University and met with over a thousand students, faculty, and staff using ...

googleforstudents.blogspot.com/2008/09/google-apps-bus-stops-at-beginning.html

[More results from Google Students »](#)

Searches related to: **google bus**

[tamil nadu bus](#)

[google apps bus](#)

[google bus routes](#)

[google bus transit](#)

[Results from your social circle for seattle](#) - BETA - [My social circle](#) - [My social content](#)

1078 photos - 17 contacts - Last photo 3 months ago



[Results from people in your social circle for san francisco international airport hotel](#) - BETA - [My social circle](#) - [My social content](#)

[San Francisco Airport Hotel Burlingame California](#)



[Crowne Plaza SFO](#) - connected via twitter.com

Our Burlingame **hotel** is only 1.5 miles south of **San Francisco International Airport** on the San Francisco Bay close to an array of exciting attractions. ...

www.sfocp.com/

[More results from Crowne Plaza SFO »](#)



Google's Social Search



News results for **jesus**



[Ha'aretz](#)

[Archbishop of Wales gives his Easter sermon at Llandaff Cathedral](#) -

2 hours ago

"But the Easter story reminds us constantly that God, through **Jesus** ... She said: "If I were to ask people on the street today 'Have you seen **Jesus** Christ?"

...

[WalesOnline](#) - [1961 related articles](#) »

[Taking Up the Dr. Seuss School of Catholicism](#) - [TIME](#) - [96 related articles](#) »

[Disturbing questions at Easter](#) - [Jamaica Gleaner](#) - [93 related articles](#) »

Latest results for **jesus** - [Pause](#)

Jer: It's gonna be 79 today!? Matt: **Jesus**?

[happyinc77](#) - [Twitter](#) - seconds ago

RT [@alaintha](#): [@kirstiealley](#) happy **jesus** resurection day

[tinytott67](#) - [Twitter](#) - seconds ago

Jesus Christ Noel, dial down the mental would you? It's Deal or No Deal, not Twin Peaks

[doubleshiny](#) - [Twitter](#) - seconds ago



Aardvark

The top part of the image shows a screenshot of the Aardvark web interface. At the top, it says "Ask a question and I'll find someone to answer". Below this is a text input field containing "What's the best...". To the left of the input field is a grey person icon. Below the input field, there is a link for "Example questions" and a prominent orange button labeled "Ask someone".

The bottom part of the image is a diagram illustrating the Aardvark workflow in three steps:

- 1. Send Aardvark a question**: An orange person icon asks, "What's a great biking path around Golden Gate Park?" (Sent 11:16 AM PDT).
- 2. Aardvark finds the perfect person to answer**: An orange path with arrows leads from the questioner to a group of diverse people icons.
- 3. Get their response in a few minutes**: A green person icon responds, "My favorite is a secret trail that takes you to the beach..." (Sent 11:20 AM PDT).



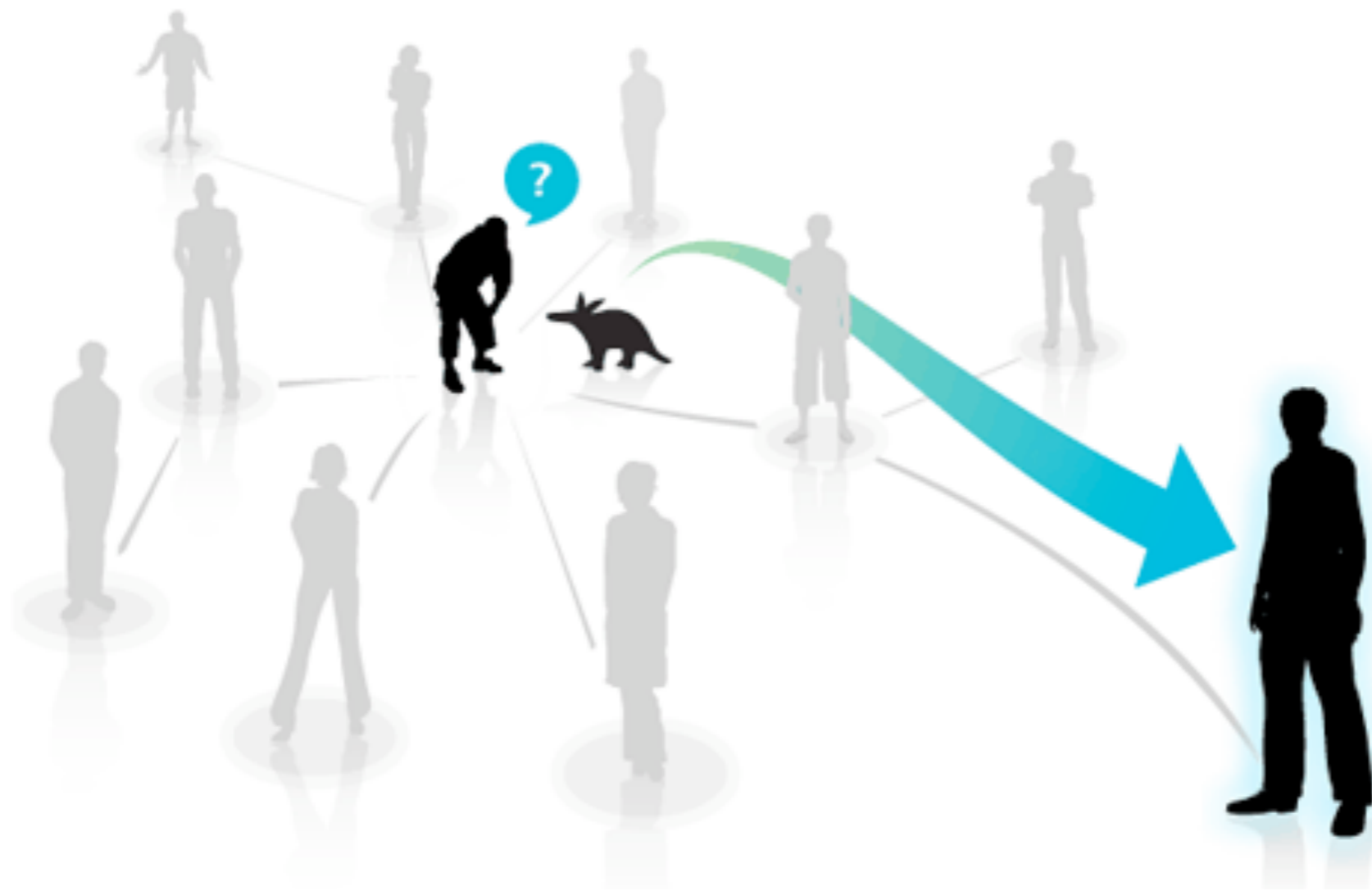
Evolution of Search

- Question
 - Contents
 - Machine Intelligence (Dialog systems)
 - People
 - Friends
 - Hybrid



The Anatomy of A Large-Scale Social Search Engine

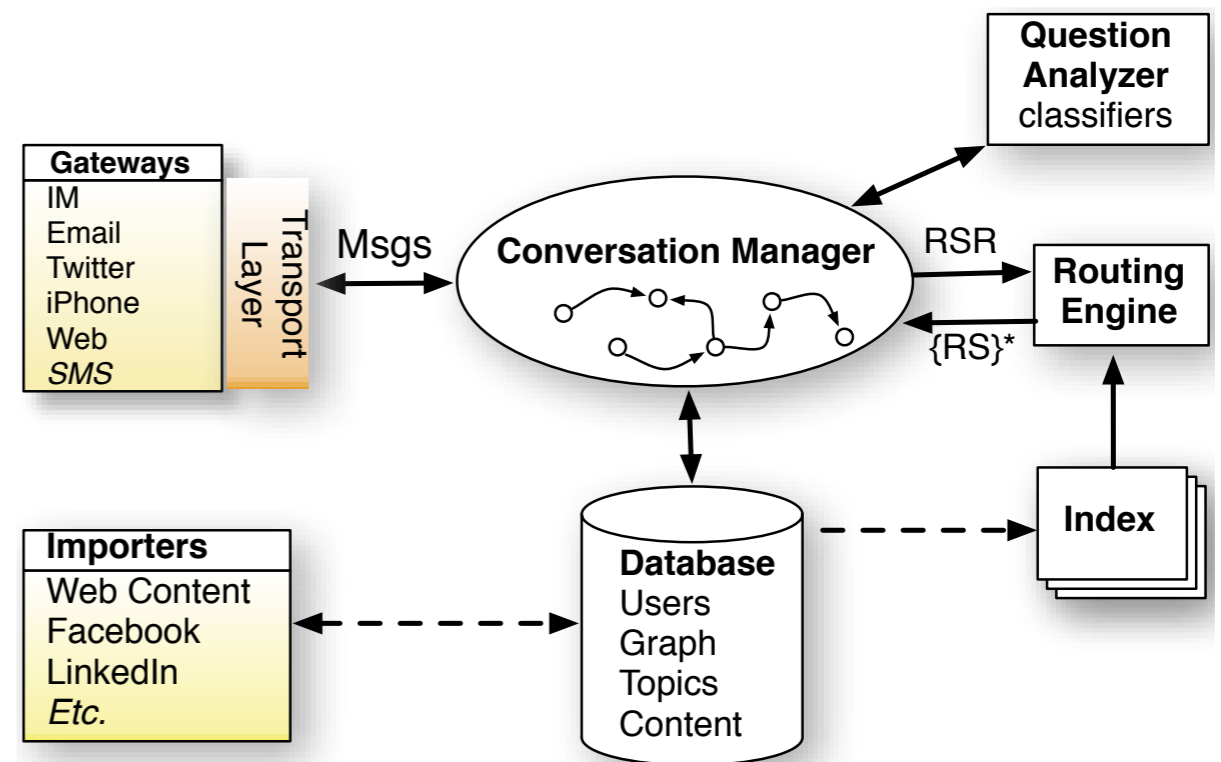
[D. Horowitz et al., WWW2010]



The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]

- Main components
 - Crawler and Indexer
 - Query Analyzer
 - Ranking Function
 - UI



The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]

- The model
 - With the topics T , the probability that user i will successfully answer question q is defined as

$$p(u_i|q) = \sum_{t \in T} p(u_i|t)p(t|q)$$

- Given a question q from user j , return a ranked list of user i that maximizes $s(u_i, u_j, q)$

$$s(u_i, u_j, q) = p(u_i|u_j) \cdot p(u_i|q) = p(u_i|u_j) \sum_{t \in T} p(u_i|t)p(t|q)$$



The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]

aardvark: You there? I have a question about **tennis** that I think you might be able to answer. (Type 'sure', 'pass', or 'busy'.)

me: sure

aardvark: (From Michael B./29/M/NewYork,US)
Does anyone know which health clubs in Manhattan have tennis courts? Or are all the good options in the other boroughs?
(Type 'pass' to skip, or 'more' for extra options.)

me: why

aardvark: I thought of you for Michael's question because

- Michael is a friend of your friend Sally Johnson.
- You know a lot about **racquet sports** and **fitness**
- You and he have similar interests in **sports**

(Type 'more' or change settings at <http://vark.com/a/XXXXX>)

me: Well there is always the Midtown Tennis Club on 8th ave @27th if you really want to stay in manhattan -- but the quality isn't great. You'd do just as well to use the public courts in Central Park. Or another good option is to join NYHRC or NYSC in manhattan, and use their courts in other boroughs...

aardvark: Great -- I've sent that to Michael. Thanks for the fast answer! (Type 'Michael:' followed by a message to add something, or 'more' for options.)

Figure 3: Example of Aardvark interacting with an answerer

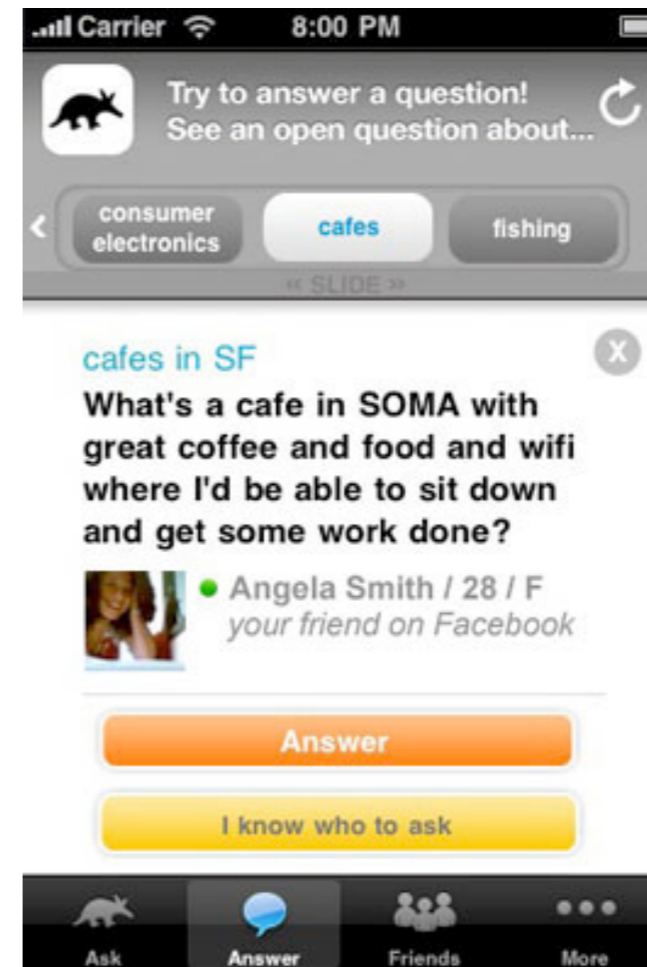


Figure 4: Screenshot of Aardvark Answering Tab on iPhone



The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]

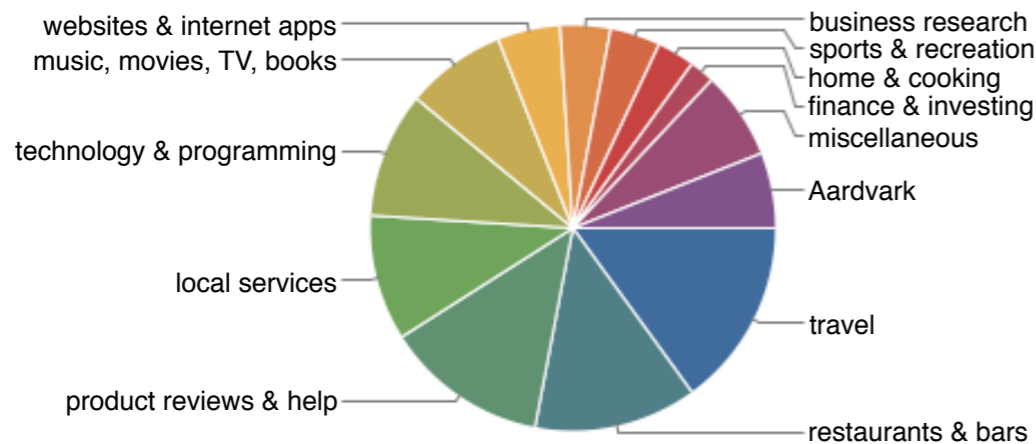


Figure 8: Categories of questions sent to Aardvark

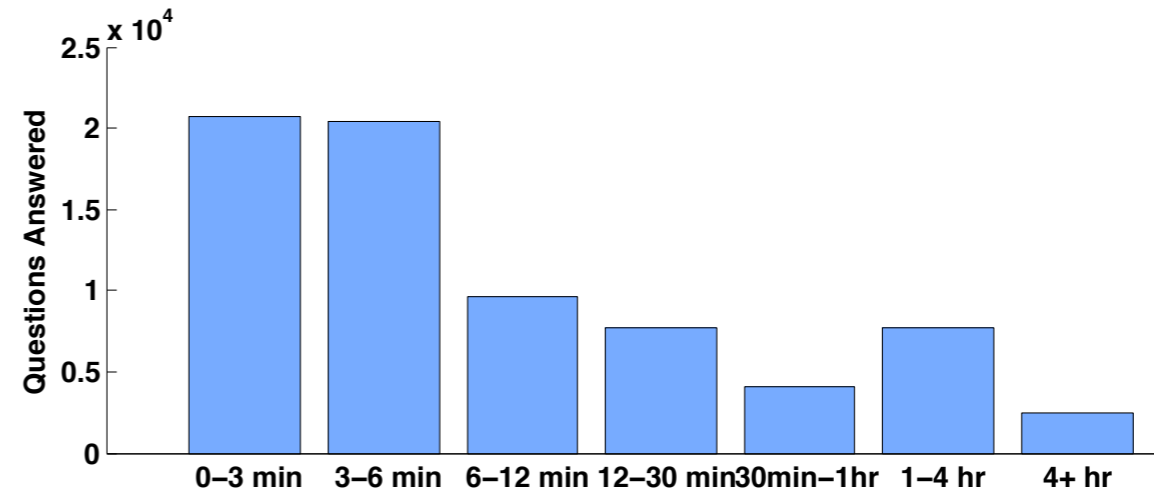


Figure 9: Distribution of questions and answering times.

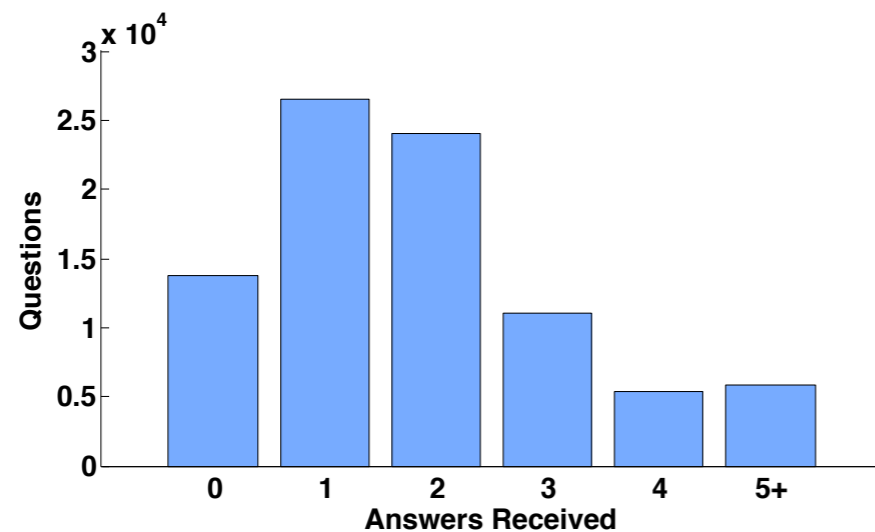


Figure 10: Distribution of questions and number of answers received.

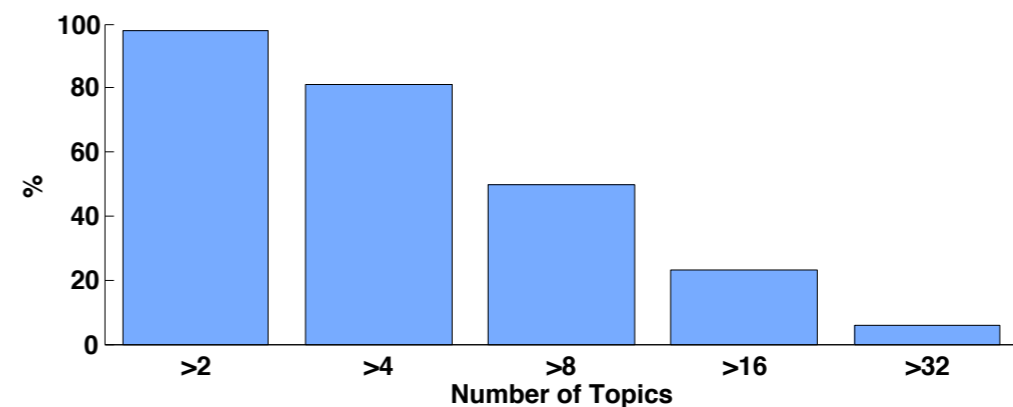


Figure 11: Distribution of percentage of users and number of topics



References

- Introducing Google Social Search: I finally found my friend's New York blog! <http://googleblog.blogspot.com/2009/10/introducing-google-social-search-i.html>
- Search Is Getting More Social. <http://googleblog.blogspot.com/2010/01/search-is-getting-more-social.html>
- D. Horowitz, S. D. Kamvar. The Anatomy of a Large Scale Social Search Engine. WWW, 2010



Outline

- Social Search Engine
- Social Recommender Systems
- Social Media Analysis



Social Recommender Systems

- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- Social-based Recommender Systems



How Much Information Is on the Web?

flickr™



amazon.com.



You Tube



ebay

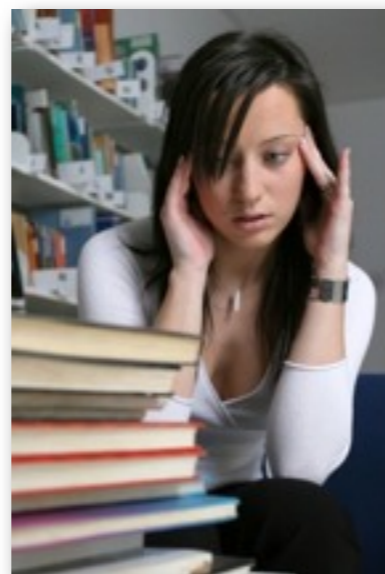
facebook

hulu™

twitter



Information Overload



Real Life Examples

Amazon.com: Social Computing, Behavioral Modeling, and Prediction: Huan Liu, John J. Salerno, Michael J. Young: Books

http://www.amazon.com/Social-Computing-Behavioral-Modeling-Prediction/ amazon

Apple Yahoo! Google Maps YouTube Wikipedia News (26) Popular

Amazon.com: Social Comp...

Click to **LOOK INSIDE!**

Social Computing, Behavioral Modeling, and Prediction (Hardcover)
by [Huan Liu](#) (Editor), [John J. Salerno](#) (Editor), [Michael J. Young](#) (Editor)
Key Phrases: [social network analysis](#), [electronic institutions](#), [cognitive modeling](#), [New York](#), [Cambridge University Press](#), [Virtual World](#) (more...)
No customer reviews yet. [Be the first.](#)

List Price: \$129.00
Price: **\$129.00** & this item ships for **FREE with Super Saver Shipping.** [Details](#)

In Stock.
Ships from and sold by **Amazon.com**. Gift-wrap available.

Want it delivered **Thursday, February 19?** Order it in the next **14 hours and 52 minutes**, and choose **One-Day Shipping** at checkout. [Details](#)

12 new from \$95.29 **4 used** from \$103.33

[Share your own customer images](#)
[Search inside this book](#)

Tell the Publisher!
[I'd like to read this book on Kindle](#)
Don't have a Kindle? [Get yours here.](#)

Quantity: 1

[Add to Shopping Cart](#)
or
[Sign in](#) to turn on 1-Click ordering
or
[Add to Cart with FREE Two-Day Shipping](#)
Amazon Prime Free Trial required. Sign up when you check out. [Learn More](#)

More Buying Choices
16 used & new from \$95.29
Have one to sell? [Sell yours here](#)

[Add to Wish List](#) | ▾
[Add to Shopping List](#)
[Add to Wedding Registry](#)
[Add to Baby Registry](#)
[Share with Friends](#)

Frequently Bought Together
Customers buy this book with [Social Network Analysis: A Handbook](#) by John P Scott
Price For Both: \$174.85



Real Life Examples

Amazon.com: Social Computing, Behavioral Modeling, and Prediction: Huan Liu, John J. Salerno, Michael J. Young: Books

http://www.amazon.com/Social-Computing-Behavioral-Modeling-Prediction/

Apple Yahoo! Google Maps YouTube Wikipedia News (26) Popular

Amazon.com: Social Comp...

Frequently Bought Together

Customers buy this book with [Social Network Analysis: A Handbook](#) by John P Scott

LOOK INSIDE! + LOOK INSIDE!

Price For Both: **\$174.85**

Add both to Cart

Add to wedding registry

Add to Baby Registry

Share with Friend

Customers Who Bought This Item Also Bought

Page

LOOK INSIDE! **PREDICTABLY IRRATIONAL: THE POWER OF NEUTRALITY** by Daniel Ariely

LOOK INSIDE! **COMPLEX**

LOOK INSIDE! **Generative Social Science: Studies in Agent-Bas...** by Joshua M. Epstein

★★★★☆ (18) \$16.34

★★★★★ (5) \$42.00

Five scales rating

★ I hate it

★★ I don't like it

★★★ It's ok

★★★★ I like it

★★★★★ I love it



Editorial Reviews

Product Description

Social computing concerns... reproduces the social behavior, and allows for experimenting with and deep understanding of behavior, patterns, and potential



Real Life Examples

amazon.com | Hao's Amazon.com | See All 40 Product Categories | Your Account |  Cart | Your Lists | Help | 

Your Browsing History | Recommended For You | Rate These Items | Improve Your Recommendations | Your Profile | Learn More

Today's Recommendations For You

Here's a daily sample of items recommended for you. Click here to [see all recommendations](#).

Page 1 of 25



[Invincible](#) ~ Michael Jackson
★★★★☆ (880) \$7.99



[In Search of Sunrise, Vol. 7: Asia](#) ~ DJ Tiesto
★★★★☆ (53) \$15.99



[Fallen](#) ~ Evanescence
★★★★☆ (2,447) \$8.99



[Amar Es Combatir](#) ~ Maná
★★★★☆ (55) \$8.49



Real Life Examples

YAHOO! MOVIES

My Movies: gabe_ma [Edit Profile](#)

Recommendations For You

[Receive Recommendations by Email](#)

Movies in Theaters: 94089



Burn After Reading (R)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **B-** 4794 ratings

The Critics: **B** 14 reviews

Don't Recommend Again Seen It? Rate It!



Pride and Glory (R)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **A-** 59 ratings

The Critics: **C+** 6 reviews

Don't Recommend Again Seen It? Rate It!



Fight Club (R)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **B+** 52392 ratings

The Critics: **B** 12 reviews

Don't Recommend Again Seen It? Rate It!



Lakeview Terrace (PG-13)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **B** 3229 ratings

The Critics: **C** 12 reviews

Don't Recommend Again Seen It? Rate It!



Vicky Cristina Barcelona (PG-13)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **B** 1923 ratings

The Critics: **B+** 13 reviews

Don't Recommend Again Seen It? Rate It!



The Duchess (PG-13)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **B+** 953 ratings

The Critics: **B-** 10 reviews

Don't Recommend Again Seen It? Rate It!

[See All Recommendations](#)





Real Life Examples



Introduction to Social Computing, Irwin King, 2010 ELL PhD School: Cloud Computing, Service Computing & Social Networks, November 23-27, 2010, Brisbane, Australia



Real Life Examples

iLike™

Songs from friends and similar people

[▶ Play All](#)  | [Buy all](#) 



[▶ Victims by The Oppressed](#)

New! Traditional Byrd69



[▶ Skinhead Girl by The Oppressed](#)

New! Traditional Byrd69



[▶ King Of The Jungle by Last Resort](#)

New! Traditional Byrd69



[▶ Violence In Our Minds by Last Resort](#)

New! Traditional Byrd69



[▶ Violence by The Templars](#)

New! Traditional Byrd69



[View all](#) | [invite more friends](#)



Basic Approaches

- Content-based Filtering
 - Recommend items based on **key-words**
 - More appropriate for information retrieval
- **Collaborative Filtering (CF)**
 - Look at users with similar rating styles
 - Look at similar items for each item

Underling assumption: personal tastes are correlated--
Active user will prefer those items which the similar users prefer.



Framework

		Items											
		i_1	i_2			i_j						i_m	
Users	u_1												
	u_2	1	3		4		2		5			3	4
	u_i		3		4		r_{ij}	3	4		3	4	4
	u_n	1			3	5	2		4	1			3

- The tasks

- Find the unknown rating?
- Which item should be recommended?



Social Recommender Systems

- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- Social-based Recommender Systems

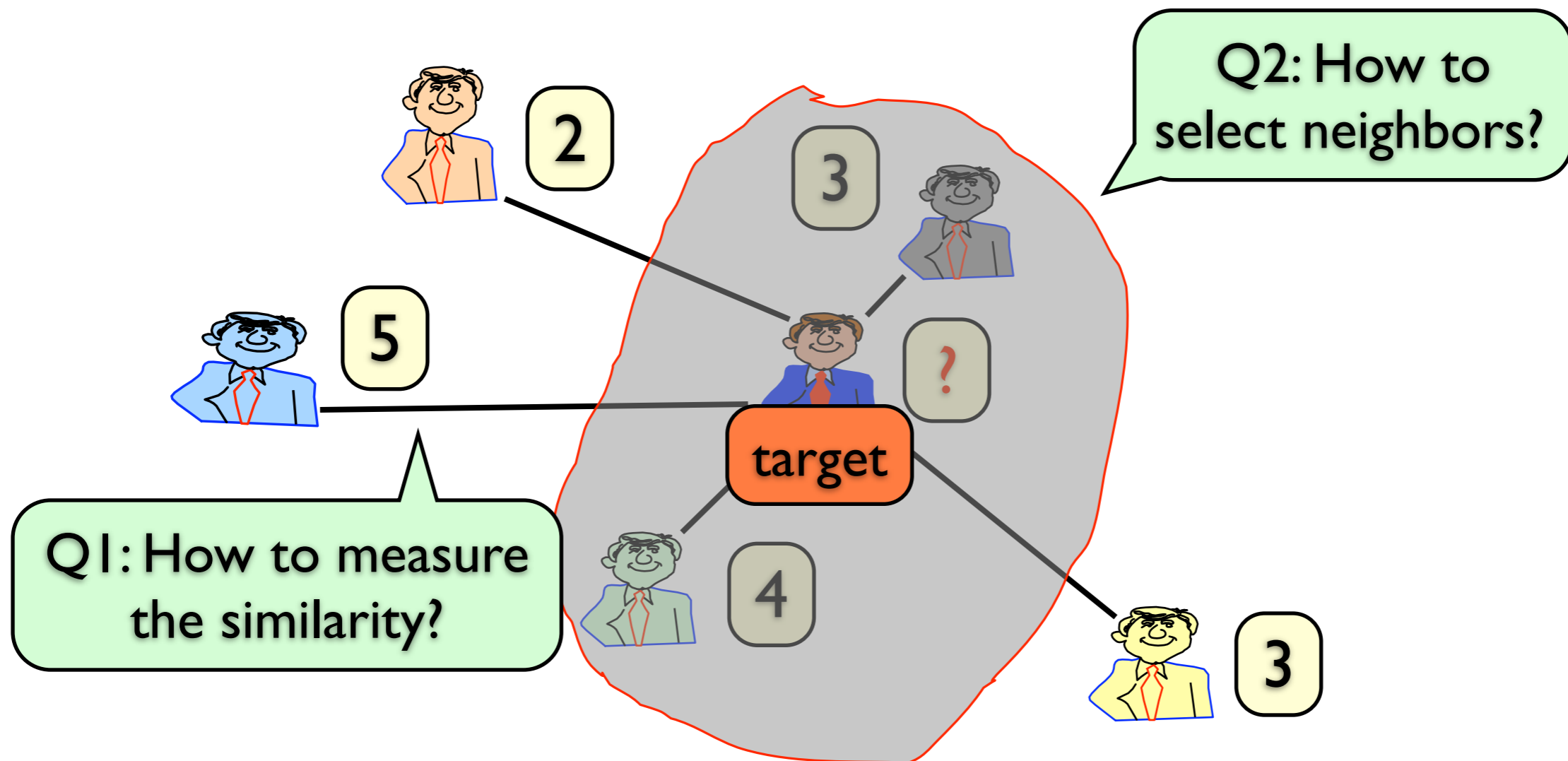


Collaborative Filtering

- Memory-based (Neighborhood-based)
 - User-based
 - Item-based
- Model-based
 - Clustering Methods
 - Bayesian Methods
 - Matrix Factorization
 - etc.



User-User Similarity



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3		4		2		5			3	4	
u ₃													
u ₄		3		4			3	4		3	4		4
u ₅													
u ₆	1			3	5	2		4	1			3	



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3		4		2		5			3	4	
u ₃													
u ₄		3		4			3	4		3	4		4
u ₅													
u ₆	1			3	5	2		4	1			3	



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3	4	2	5			3	4				
u ₃													
u ₄		3	4		3	4		3	4		4		
u ₅													
u ₆	1		3	5	2	4	1				3		



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3		4		2		5			3	4	
u ₃													
u ₄		3		4			3	4		3	4		4
u ₅													
u ₆	1			3	5	2		4	1			3	



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3	4		2		5			3	4		
u ₃													
u ₄		3	4			3	4		3	4		4	
u ₅													
u ₆	1		3	5	2		4	1			3		



User-based Collaborative Filtering

- Predict the ratings of active users based on the ratings of similar users found in the user-item matrix
- Pearson correlation coefficient

$$w(a, i) = \frac{\sum_j (r_{aj} - \bar{r}_a)(r_{ij} - \bar{r}_i)}{\sqrt{\sum_j (r_{aj} - \bar{r}_a)^2 \sum_j (r_{ij} - \bar{r}_i)^2}} \quad j \in I(a) \cap I(i)$$

- Cosine measure

$$c(a, i) = \frac{r_a \cdot r_i}{\|r_a\|_2 * \|r_i\|_2}$$

u_i	1	3	4	2		5		3	4		
u_a		3	4			3	4	3	4	4	
	1		3	5	2		4	1		3	



Collaborative Filtering

- Memory-based (Neighborhood-based)
 - User-based
 - Item-based
- Model-based
 - Clustering Methods
 - Bayesian Methods
 - Matrix Factorization
 - etc.



Item-Item Similarity

- Search for similarities among items
- Item-Item similarity is more stable than user-user similarity



Correlation-based Methods

[Sarwar, 2001]

- Same as in user-user similarity but on item vectors
- Pearson correlation coefficient
 - Look for users who rated both items

$$s_{ij} = \frac{\sum_u (r_{uj} - \bar{r}_j)(r_{ui} - \bar{r}_i)}{\sqrt{\sum_u (r_{uj} - \bar{r}_j)^2 \sum_u (r_{ui} - \bar{r}_i)^2}}$$

- u : users rated both items

	i_1	i_2			i_i	i_j				i_m
u_1										
u_2	1	3		4	2	5			3	4
u_i		3		4		3	4		3	4
u_n	1			3	5	2	4	1		3



Collaborative Filtering

- Memory-based (Neighborhood-based)
 - User-based
 - Item-based
- Model-based
 - Clustering Methods
 - Bayesian Methods
 - **Matrix Factorization**
 - etc...



Matrix Factorization

	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2		3		4		
u_2	4	3			5			
u_3	4		2				2	4
u_4								
u_5	5	1	2		4	3		
u_6	4	3		2	4		3	5

	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2	2.5	3	4.8	4	2.2	4.8
u_2	4	3	2.4	2.9	5	4.1	2.6	4.7
u_3	4	1.7	2	3.2	3.9	3.0	2	4
u_4	4.8	2.1	2.7	2.6	4.7	3.8	2.4	4.9
u_5	5	1	2	3.4	4	3	1.5	4.6
u_6	4	3	2.9	2	4	3.4	3	5

$$U = \begin{bmatrix} 1.55 & 1.22 & 0.37 & 0.81 & 0.62 & -0.01 \\ 0.36 & 0.91 & 1.21 & 0.39 & 1.10 & 0.25 \\ 0.59 & 0.20 & 0.14 & 0.83 & 0.27 & 1.51 \\ 0.39 & 1.33 & -0.43 & 0.70 & -0.90 & 0.68 \\ 1.05 & 0.11 & 0.17 & 1.18 & 1.81 & 0.40 \end{bmatrix}$$

$$V = \begin{bmatrix} 1.00 & -0.05 & -0.24 & 0.26 & 1.28 & 0.54 & -0.31 & 0.52 \\ 0.19 & -0.86 & -0.72 & 0.05 & 0.68 & 0.02 & -0.61 & 0.70 \\ 0.49 & 0.09 & -0.05 & -0.62 & 0.12 & 0.08 & 0.02 & 1.60 \\ -0.40 & 0.70 & 0.27 & -0.27 & 0.99 & 0.44 & 0.39 & 0.74 \\ 1.49 & -1.00 & 0.06 & 0.05 & 0.23 & 0.01 & -0.36 & 0.80 \end{bmatrix}$$



Matrix Factorization

- Matrix Factorization in Collaborative Filtering
 - To fit the product of two (low rank) matrices to the observed rating matrix.
 - To find two latent user and item feature matrices.
 - To use the fitted matrix to predict the unobserved ratings.

$$\begin{pmatrix} \mathbf{u}_{11} & \cdots & \mathbf{u}_{1k} \\ \vdots & \ddots & \vdots \\ \mathbf{u}_{m1} & \cdots & \mathbf{u}_{mk} \end{pmatrix} \begin{pmatrix} \mathbf{v}_{11} & \cdots & \mathbf{v}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{v}_{k1} & \cdots & \mathbf{v}_{kn} \end{pmatrix}$$

User-specific latent feature vector

Item-specific latent feature column vector



Matrix Factorization

- Optimization Problem
- Given a $m \times n$ rating matrix R , to find two matrices $U \in \mathbb{R}^{l \times m}$ and $V \in \mathbb{R}^{l \times n}$,

$$R \approx U^T V,$$

where $l < \min(m, n)$, is the number of factors



Matrix Factorization

- Models
 - SVD-like Algorithm
 - Regularized Matrix Factorization (RMF)
 - Probabilistic Matrix Factorization (PMF)
 - Non-negative Matrix Factorization (NMF)



SVD-like Algorithm

- Minimizing

$$\frac{1}{2} \|R - U^T V\|_F^2,$$

- For collaborative filtering

$$\min_{U, V} \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij} (R_{ij} - U_i^T V_j)^2$$

where I_{ij} is the indicator function that is equal to 1 if user u_i rated item v_j and equal to 0 otherwise.



Regularized Matrix Factorization

- Minimize the loss based on the observed ratings with regularization terms to avoid over-fitting problem

$$\min_{U,V} \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij} (R_{ij} - U_i^T V_j)^2 + \frac{\lambda_1}{2} \|U\|_F^2 + \frac{\lambda_2}{2} \|V\|_F^2$$

Regularization terms

where $\lambda_1, \lambda_2 > 0$.

- The problem can be solved by simple gradient descent algorithm.



Regularized Matrix Factorization

- Algorithm for RMF
 - Not convex & local optimal
 - Gradient-descent algorithm
 - Gradient computation with randomly initialized U and V

$$\frac{\partial L}{\partial u_{il}} = \lambda u_{il} - \sum_{j|(i,j) \in S} (y_{ij} - \hat{y}_{ij}) v_{jl}$$

$$\frac{\partial L}{\partial v_{il}} = \lambda v_{il} - \sum_{j|(i,j) \in S} (y_{ij} - \hat{y}_{ij}) u_{jl}$$

- Update U and V alternatively

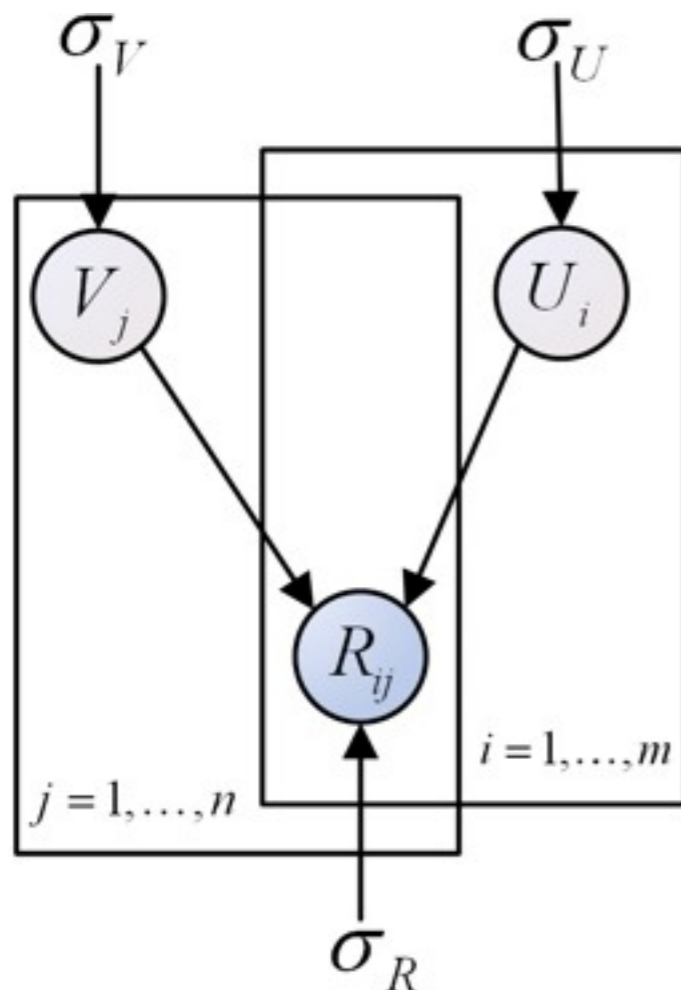
$$u_{il}^{(t+1)} = u_{il}^{(t)} - \tau \frac{\partial L}{\partial u_{il}^{(t)}}$$

$$v_{jl}^{(t+1)} = v_{jl}^{(t)} - \tau \frac{\partial L}{\partial v_{jl}^{(t)}}$$



Probabilistic Matrix Factorization

- PMF
 - Define a conditional distribution over the observed ratings as:

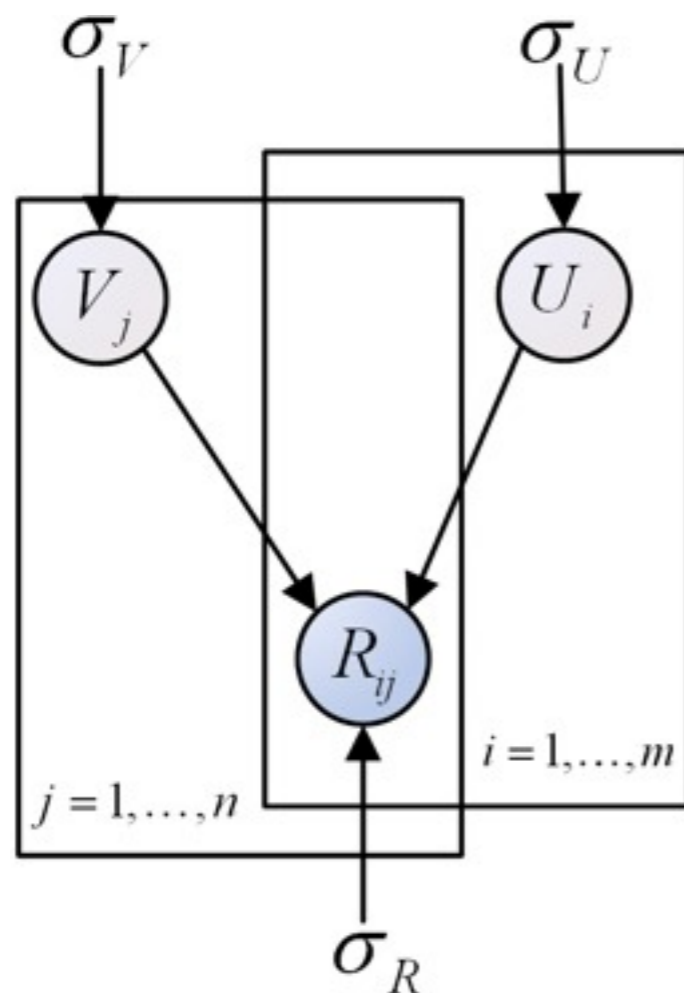


$$p(R|U, V, \sigma_R^2) = \prod_{i=1}^m \prod_{j=1}^n \left[\mathcal{N} \left(R_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right]^{I_{ij}^R}$$



Probabilistic Matrix Factorization

- PMF
 - Assume zero-mean spherical Gaussian priors on user and item feature:



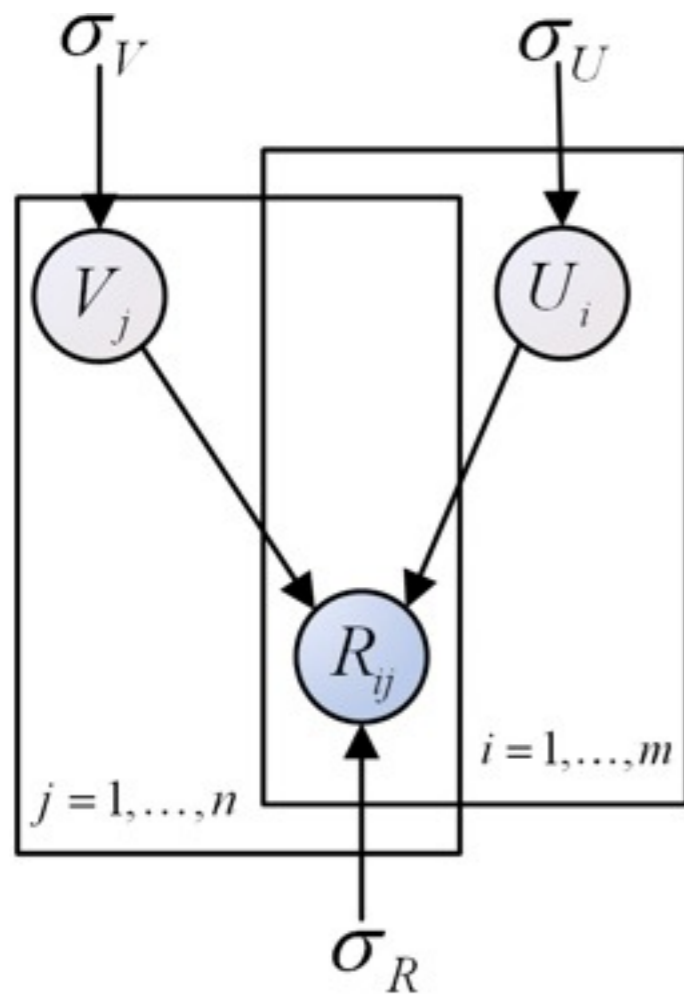
$$p(U|\sigma_U^2) = \prod_{i=1}^m \mathcal{N}(U_i|0, \sigma_U^2 \mathbf{I})$$

$$p(V|\sigma_V^2) = \prod_{j=1}^n \mathcal{N}(V_j|0, \sigma_V^2 \mathbf{I})$$



Probabilistic Matrix Factorization

- PMF
 - Bayesian inference



$$\begin{aligned} p(U, V | R, \sigma_R^2, \sigma_U^2, \sigma_V^2) &\propto p(R | U, V, \sigma_R^2) p(U | \sigma_U^2) p(V | \sigma_V^2) \\ &= \prod_{i=1}^m \prod_{j=1}^n \left[\mathcal{N} \left(R_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right]^{I_{ij}^R} \\ &\times \prod_{i=1}^m \mathcal{N}(U_i | 0, \sigma_U^2 \mathbf{I}) \times \prod_{j=1}^n \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I}). \end{aligned}$$



Non-negative Matrix Factorization

- NMF
 - Given an observed matrix Y , to find two non-negative matrices U and V
 - Two types of loss functions

- Squared error function

$$\sum_{ij} (R_{ij} - U_i^T V_j)^2$$

- Divergence

$$D(R||U^T V) = \sum_{ij} (R_{ij} \log \frac{R_{ij}}{U_i^T V_j} - R_{ij} + U_i^T V_j)$$

- Solving by multiplicative updating rules



Social Recommender Systems

- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- Social-based Recommender Systems



Challenges

- Data sparsity problem

YAHOO! MOVIES

My Movies: **gabe_ma** [Edit Profile](#)

Recommendations For You

[Receive Recommendations by Email](#)



[Watch the Trailer](#)

My Blueberry Nights (2008)

The Critics:

B-

[7 reviews](#)

My Grade:

A+

Oscar-worthy

A

B

C

D

F

Yahoo! Users:

B-

[667 ratings](#)

[write a review](#)



Vicky Cristina Barcelona (PG-13)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **B** 1923 ratings

The Critics: **B+** 13 reviews

[Don't Recommend Again](#) [Seen It? Rate It!](#)



The Duchess (PG-13)

[Showtimes & Tickets](#) | [Add to My Lists](#)

Yahoo! Users: **B+** 953 ratings

The Critics: **B-** 10 reviews

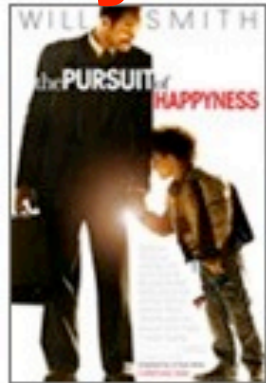
[Don't Recommend Again](#) [Seen It? Rate It!](#)

[See All Recommendations](#)



Challenges

My Movie Ratings



The Pursuit of Happyness (PG-13, 1 hr. 57 min.)

Buy DVD | Add to My Lists

Yahoo! Users: **B+** 38992 ratings

The Critics: **B-** 13 reviews

★ My Rating: A+



Finding Nemo (G, 1 hr. 40 min.)

Buy DVD | Add to My Lists

Yahoo! Users: **B+** 137394 ratings

The Critics: **A-** 14 reviews

★ My Rating: A



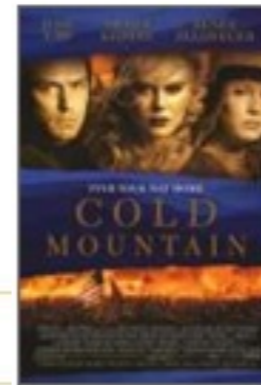
My Blueberry Nights (PG-13, 1 hr. 30 min.)

Buy DVD | Add to My Lists

Yahoo! Users: **B-** 756 ratings

The Critics: **B-** 7 reviews

★ My Rating: A+



Cold Mountain (R, 2 hrs. 35 min.)

Buy DVD | Add to My Lists

Yahoo! Users: **B** 38986 ratings

The Critics: **B+** 10 reviews

★ My Rating: B+



The Lord of the Rings: The Fellowship of the Ring

Buy DVD | Add to My Lists

Yahoo! Users: **A-** 110957 ratings

The Critics: **A** 15 reviews

★ My Rating: A



Shrek 2 (PG, 1 hr. 32 min.)

Buy DVD | Add to My Lists

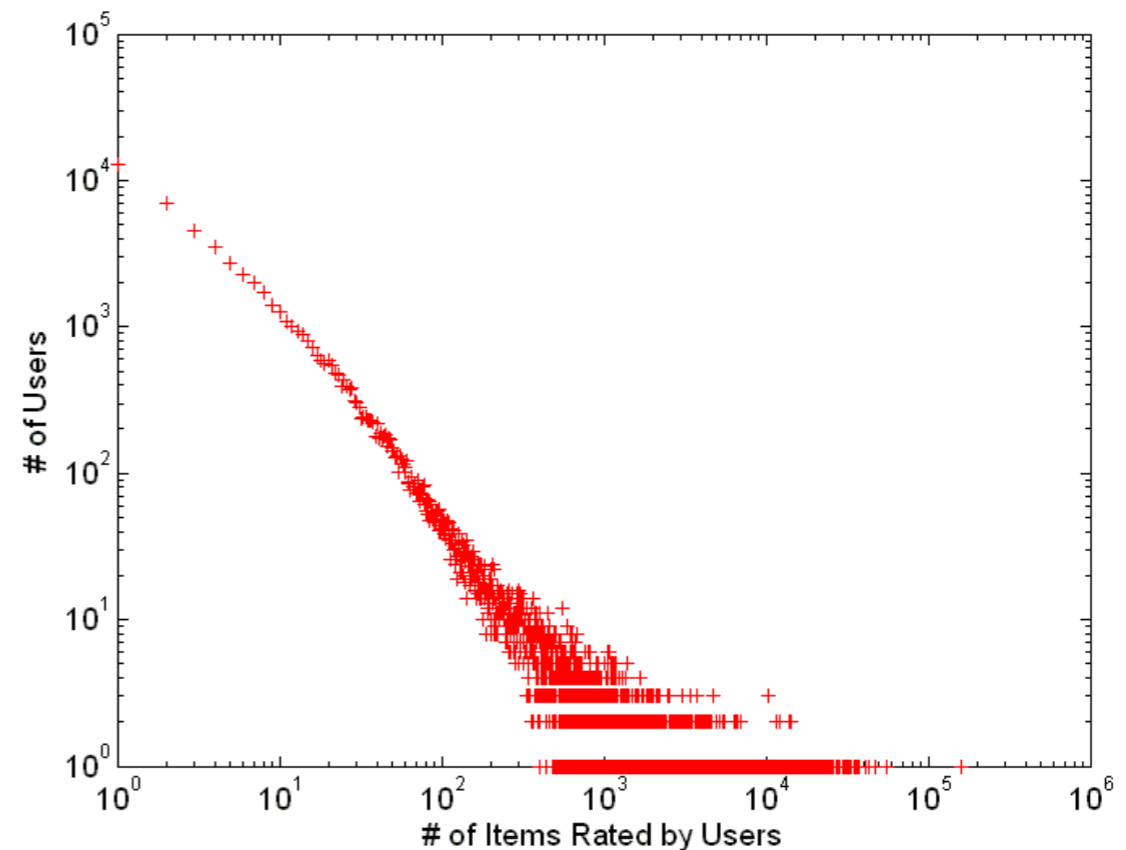
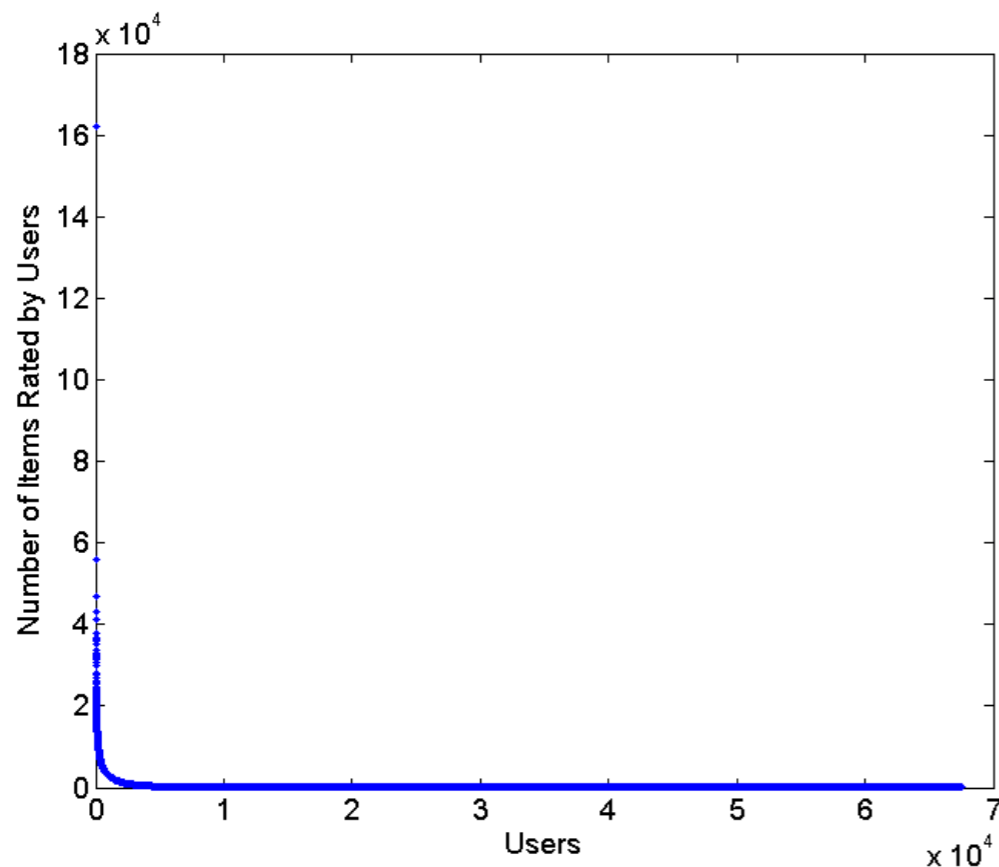
Yahoo! Users: **B+** 150368 ratings

The Critics: **B** 15 reviews

★ My Rating: B



Number of Ratings per User



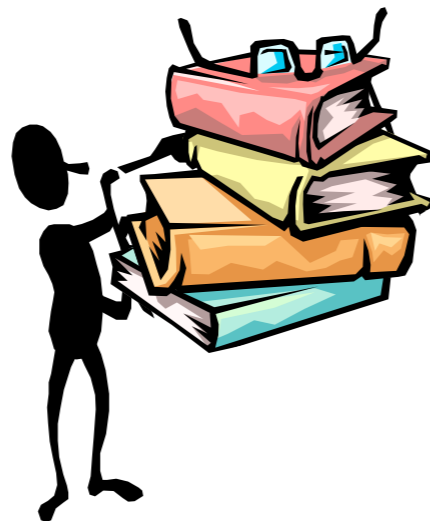
Extracted From Epinions.com

114,222 users, 754,987 items and 13,385,713 ratings



Challenges

- Traditional recommender systems ignore the social connections between users



Recommendations
from friends



Social Recommendation Using Probabilistic Matrix Factorization

[Hao Ma, et al., CIKM2008]



Motivations

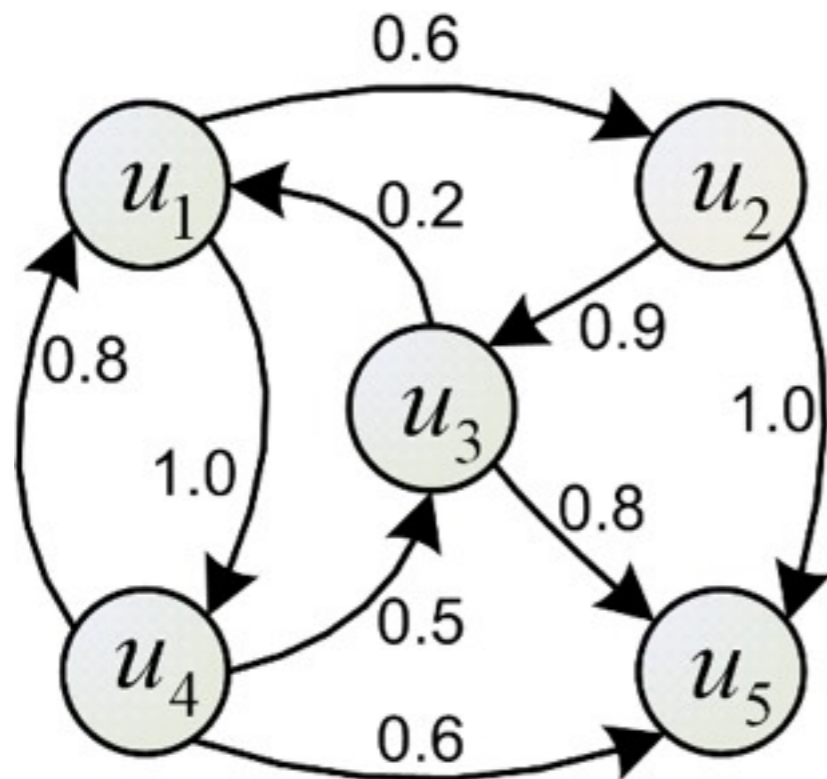
- “Yes, there is a correlation - from social networks to personal behavior on the web”

Parag Singla and Matthew Richardson ([WWW'08](#))

- Analyze the who talks to whom social network over 10 million people with their related search results
- People who chat with each other are more likely to share the same or similar interests
- To improve the recommendation accuracy and solve the data sparsity problem, **users' social network** should be taken into consideration



Problem Definition



Social Trust Graph

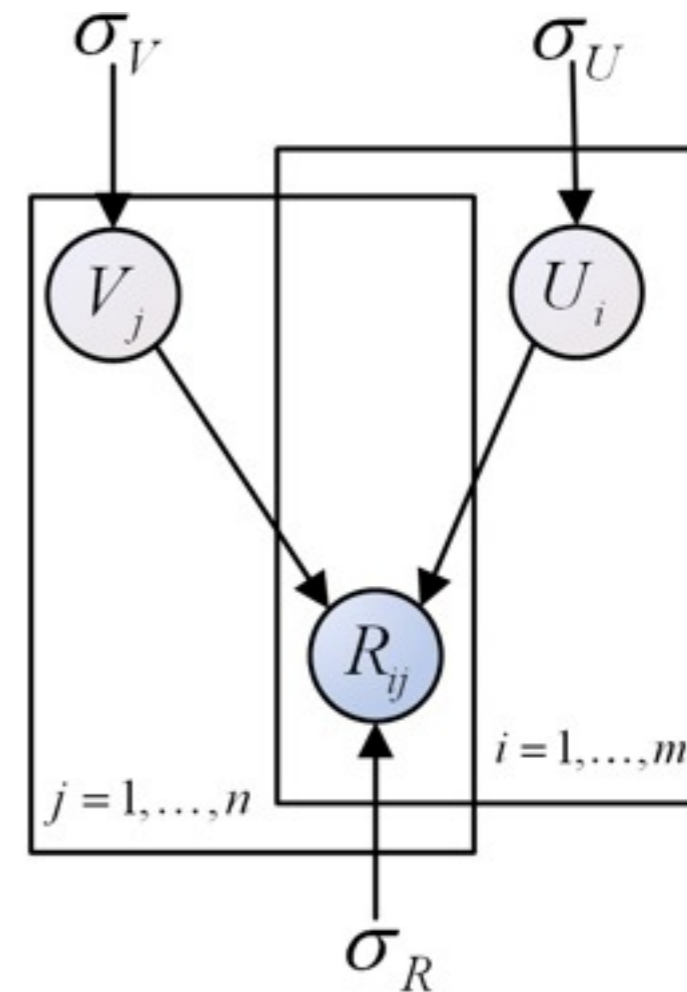
	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3

User-Item Rating Matrix



User-Item Matrix Factorization

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3



$$p(R|U, V, \sigma_R^2) = \prod_{i=1}^m \prod_{j=1}^n \left[\mathcal{N} \left(R_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right]^{I_{ij}^R}$$

$$p(U|\sigma_U^2) = \prod_{i=1}^m \mathcal{N}(U_i | 0, \sigma_U^2 \mathbf{I})$$

$$p(V|\sigma_V^2) = \prod_{j=1}^n \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I})$$

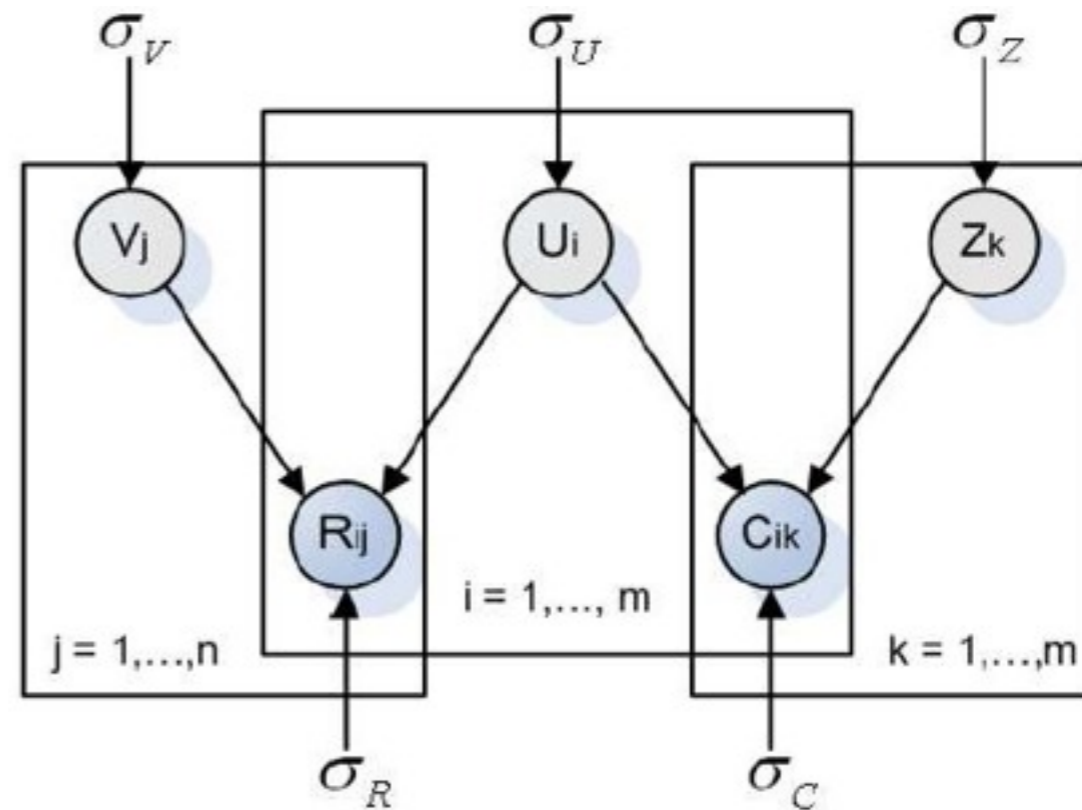
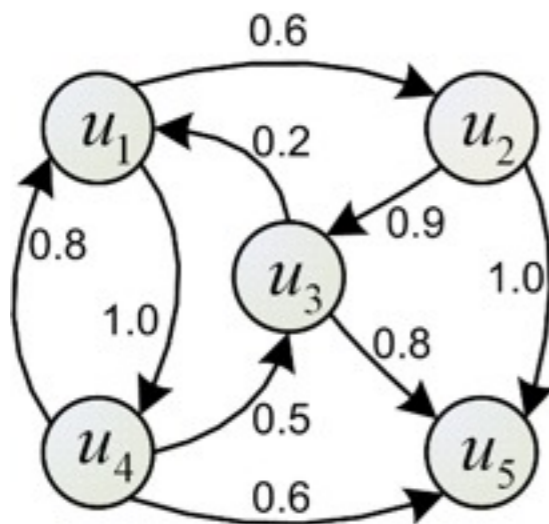
R. Salakhutdinov and A. Mnih (NIPS'08)

Introduction to Social Computing, Irwin King, 2010 ELL PhD School: Cloud Computing, Service Computing & Social Networks, November 23-27, 2010, Brisbane, Australia



SoRec

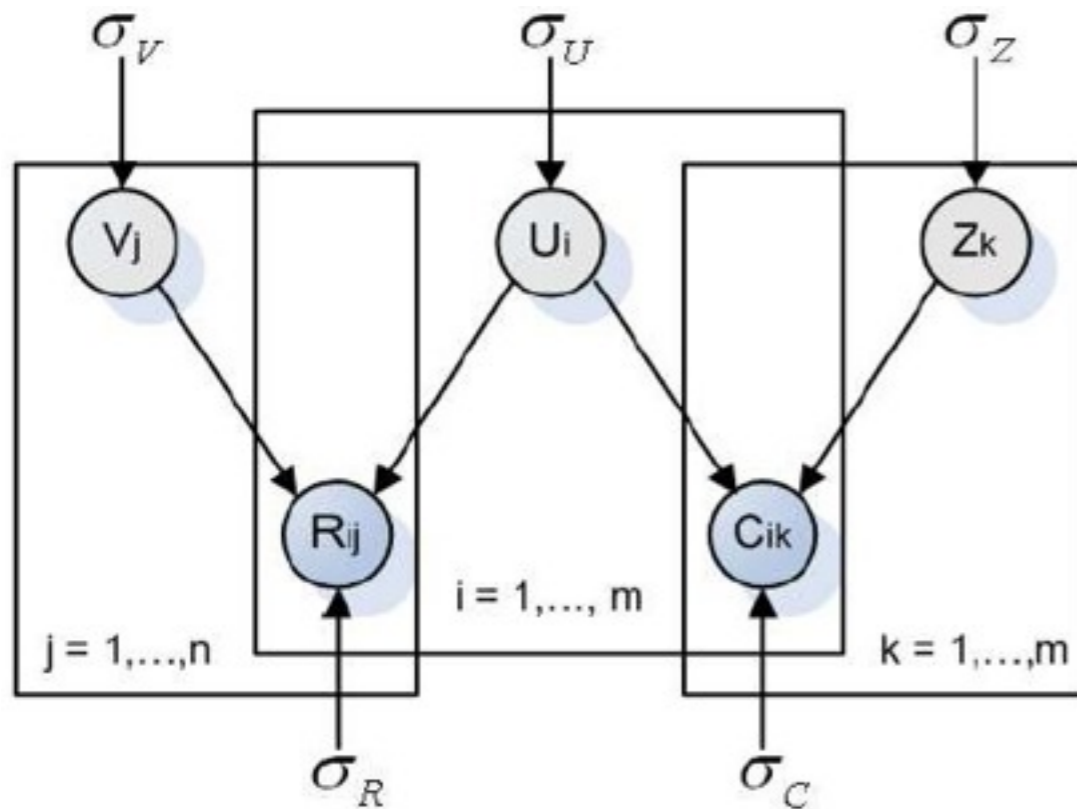
	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3



SoRec



SoRec



$$p(R|U, V, \sigma_R^2) = \prod_{i=1}^m \prod_{j=1}^n \mathcal{N} \left[\left(r_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right]^{I_{ij}^R}$$

$$p(C|U, Z, \sigma_C^2) = \prod_{i=1}^m \prod_{k=1}^m \mathcal{N} \left[\left(c_{ik} | g(U_i^T Z_k), \sigma_C^2 \right) \right]^{I_{ik}^C}$$

$$p(U|\sigma_U^2) = \prod_{i=1}^m \mathcal{N}(U_i | 0, \sigma_U^2 \mathbf{I}) \quad p(V|\sigma_V^2) = \prod_{j=1}^n \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I})$$

$$p(Z|\sigma_Z^2) = \prod_{k=1}^m \mathcal{N}(Z_k | 0, \sigma_Z^2 \mathbf{I})$$

$$\mathcal{L}(R, C, U, V, Z) =$$

$$\begin{aligned} & \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (r_{ij} - g(U_i^T V_j))^2 + \frac{\lambda_C}{2} \sum_{i=1}^m \sum_{k=1}^m I_{ik}^C (c_{ik} - g(U_i^T Z_k))^2 \\ & + \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2 + \frac{\lambda_Z}{2} \|Z\|_F^2, \end{aligned}$$



SoRec

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial U_i} &= \sum_{j=1}^n I_{ij}^R g'(U_i^T V_j) (g(U_i^T V_j) - r_{ij}) V_j \\ &+ \lambda_C \sum_{k=1}^m I_{ik}^C g'(U_i^T Z_k) (g(U_i^T Z_k) - c_{ik}^*) Z_k + \lambda_U U_i, \end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial V_j} = \sum_{i=1}^m I_{ij}^R g'(U_i^T V_j) (g(U_i^T V_j) - r_{ij}) U_i + \lambda_V V_j,$$

$$\frac{\partial \mathcal{L}}{\partial Z_k} = \lambda_C \sum_{i=1}^m I_{ik}^C g'(U_i^T Z_k) (g(U_i^T Z_k) - c_{ik}^*) U_i + \lambda_Z Z_k,$$



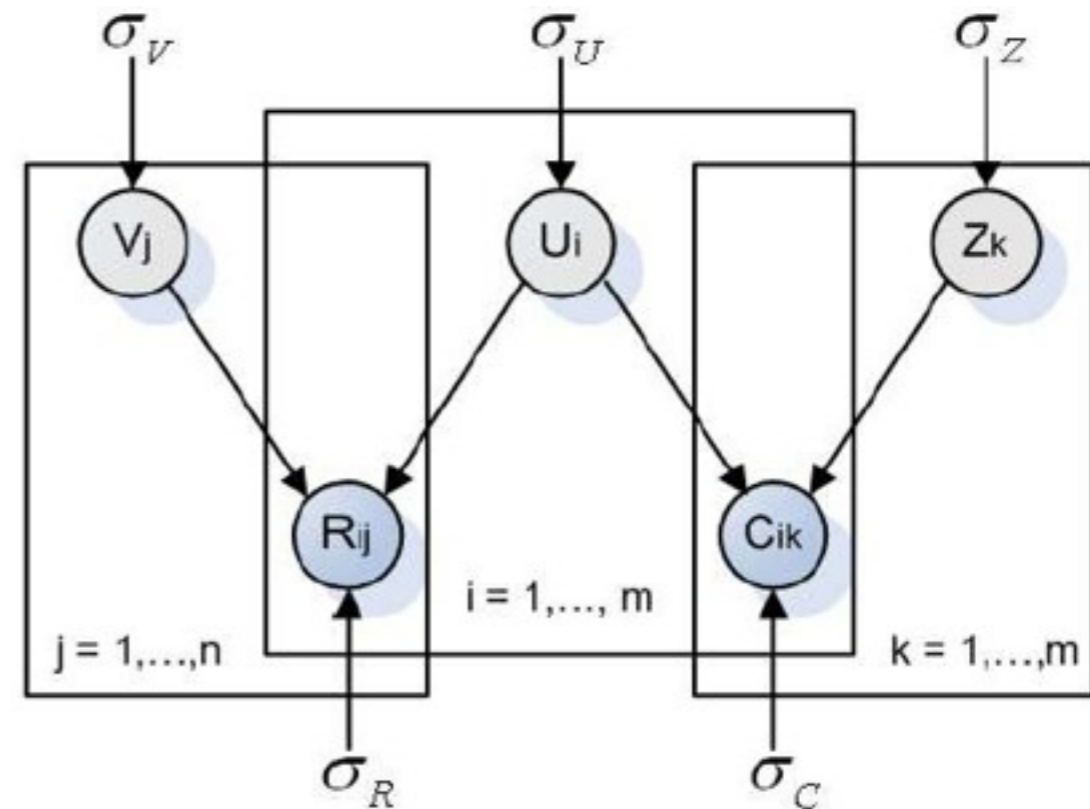
Complexity Analysis

- For the Objective Function $O(\rho_R l + \rho_C l)$
- For $\frac{\partial \mathcal{L}}{\partial U}$ the complexity is $O(\rho_R l + \rho_C l)$
- For $\frac{\partial \mathcal{L}}{\partial V}$ the complexity is $O(\rho_R l)$
- For $\frac{\partial \mathcal{L}}{\partial Z}$ the complexity is $O(\rho_C l)$
- In general, the complexity of our method is linear with the observations in these two matrices



Disadvantages of SoRec

- Lack of interpretability
- Does not reflect the real-world recommendation process



SoRec

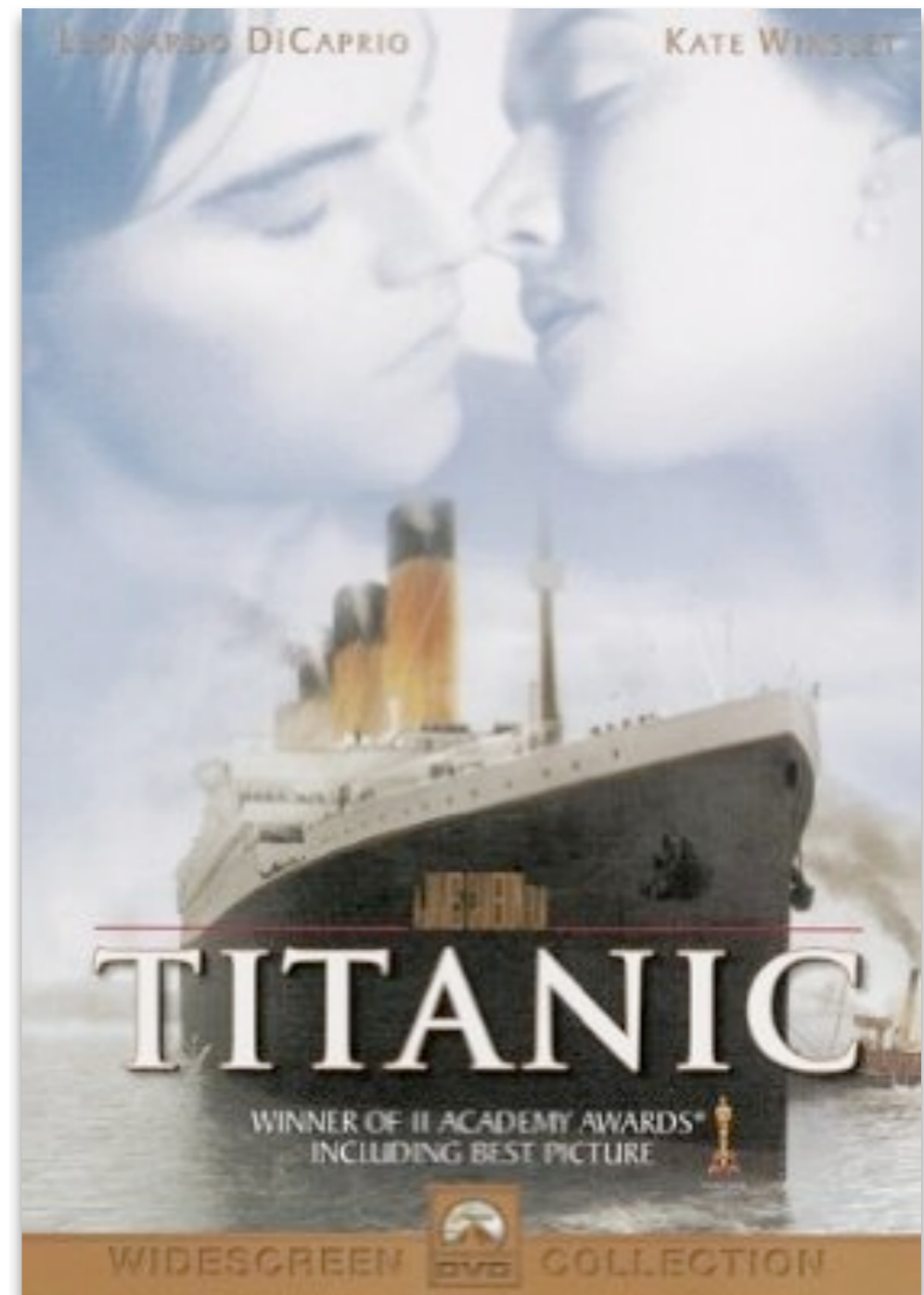


Learning to Recommend with Social Trust Ensemble

[Hao Ma, et al., SIGIR2009]



1st Motivation



1st Motivation



Introduction to Social Computing, Irwin King, 2010 EIT PhD School: Cloud Computing, Service Computing & Social Networks, November 23-27, 2010, Brisbane, Australia

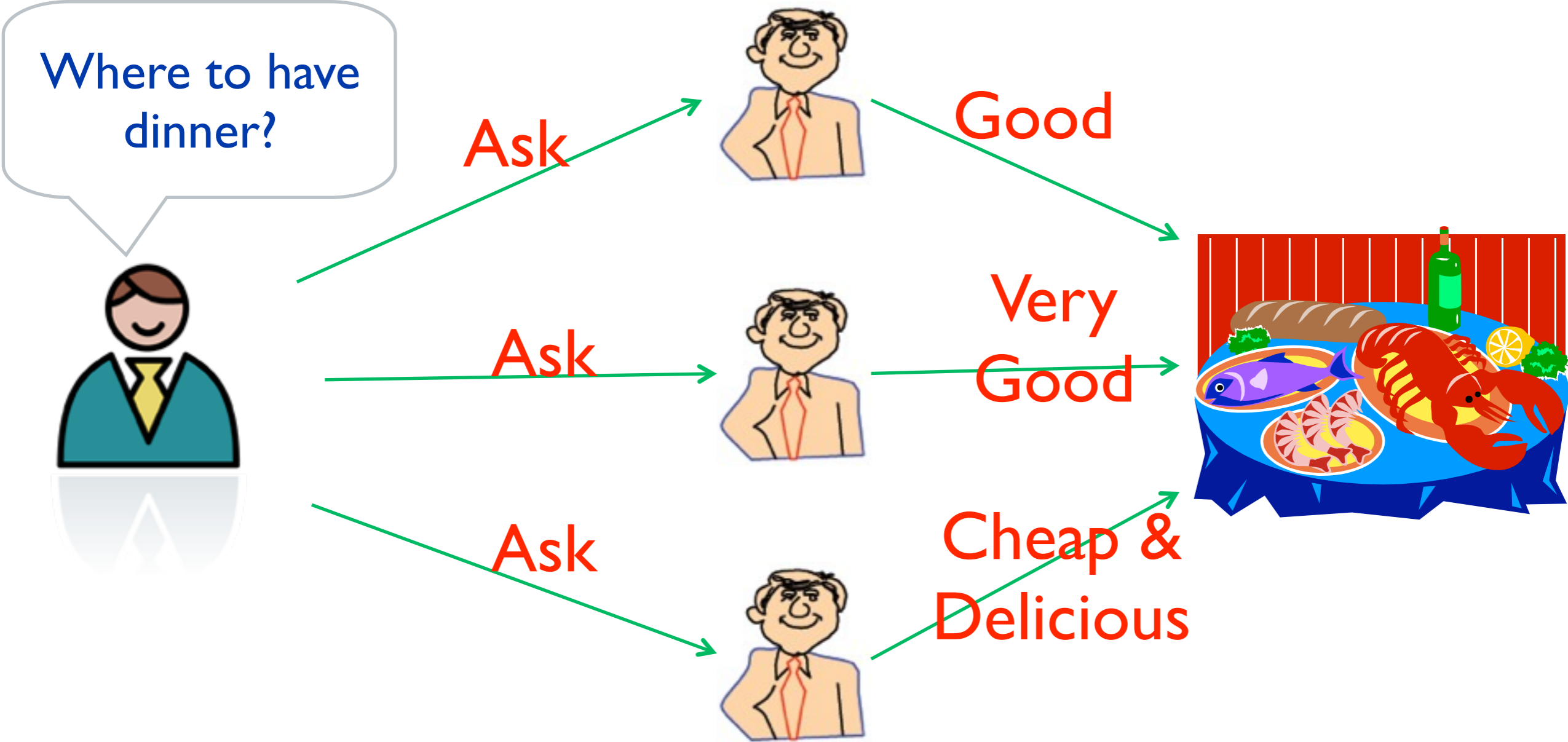


1st Motivation

- Users have their **own characteristics**, and they have different tastes on different items, such as movies, books, music, articles, food, etc.

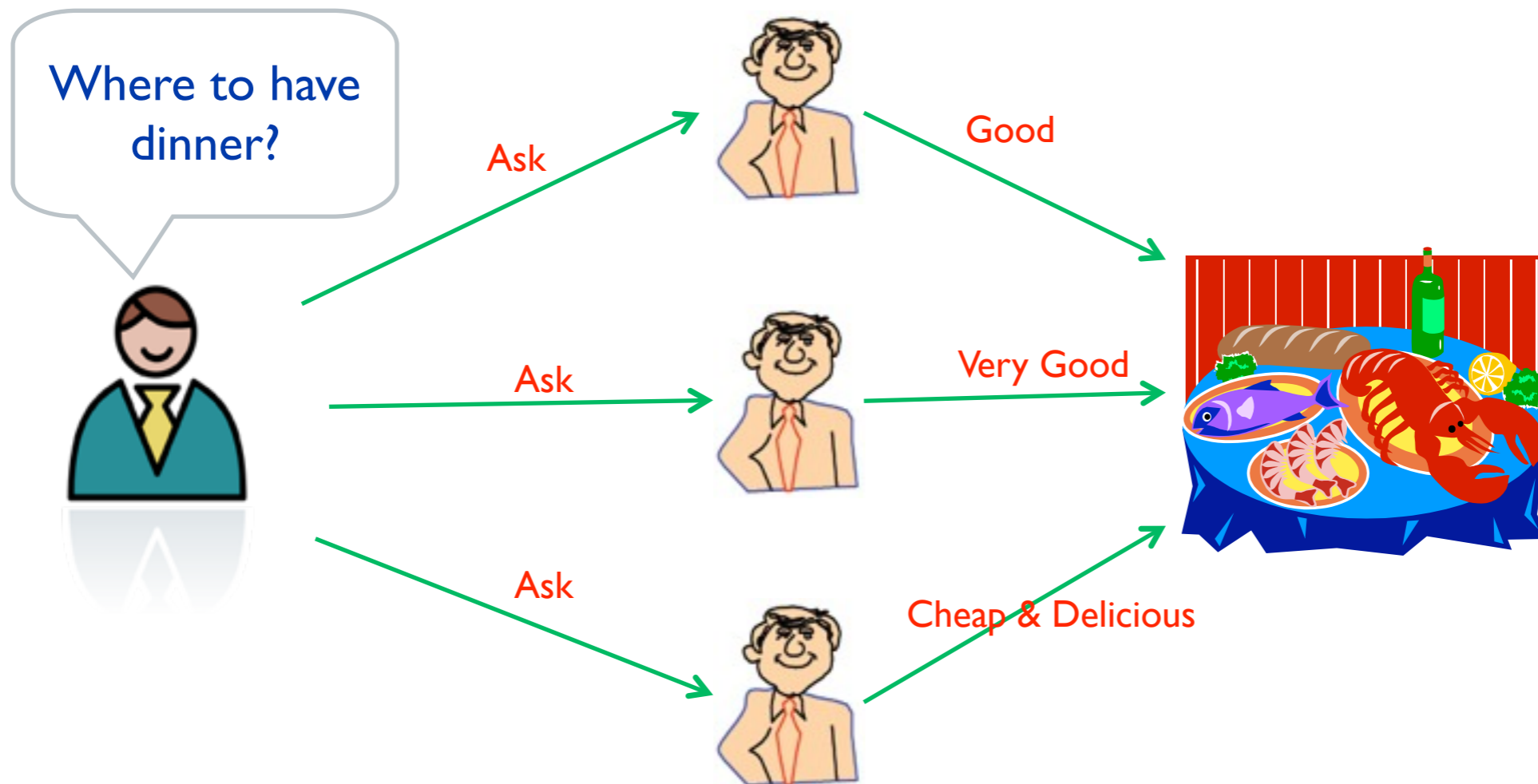


2nd Motivation



2nd Motivation

- Users can be easily **influenced by the friends they trust**, and prefer their friends' recommendations.



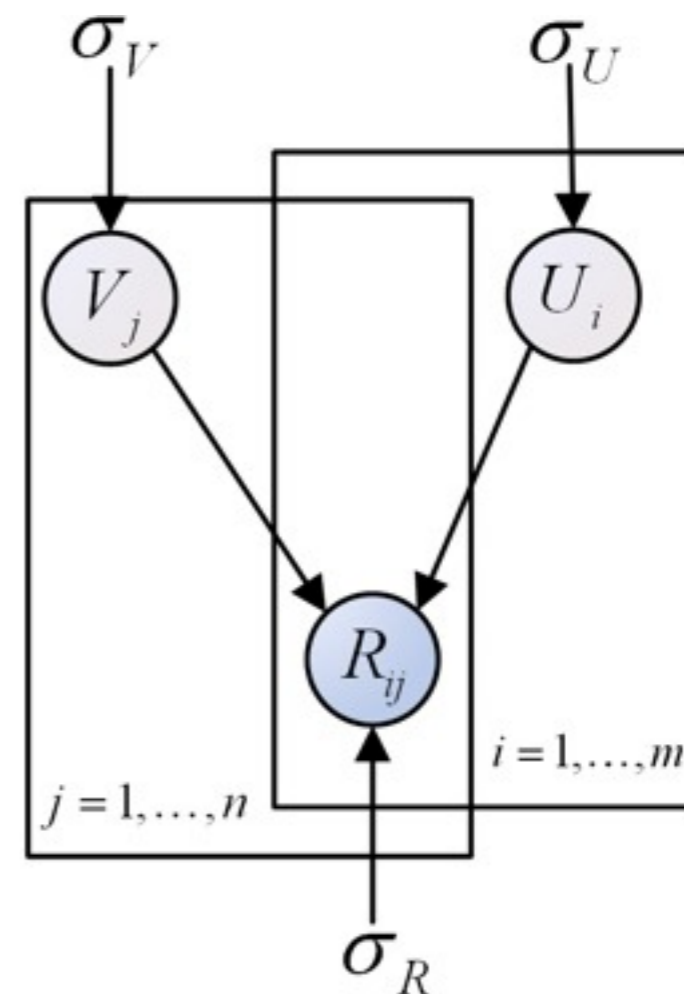
Motivations

- Users have their own characteristics, and they have different tastes on different items, such as movies, books, music, articles, food, etc.
- Users can be easily influenced by the friends they trust, and prefer their friends' recommendations.
- One user's final decision is the balance between his/her own taste and his/her trusted friends' favors.



User-Item Matrix Factorization

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3



$$p(R|U, V, \sigma_R^2) = \prod_{i=1}^m \prod_{j=1}^n \left[\mathcal{N} \left(R_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right]^{I_{ij}^R}$$

$$p(U|\sigma_U^2) = \prod_{i=1}^m \mathcal{N}(U_i | 0, \sigma_U^2 \mathbf{I})$$

$$p(V|\sigma_V^2) = \prod_{j=1}^n \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I})$$

[R. Salakhutdinov, et al., NIPS2008]



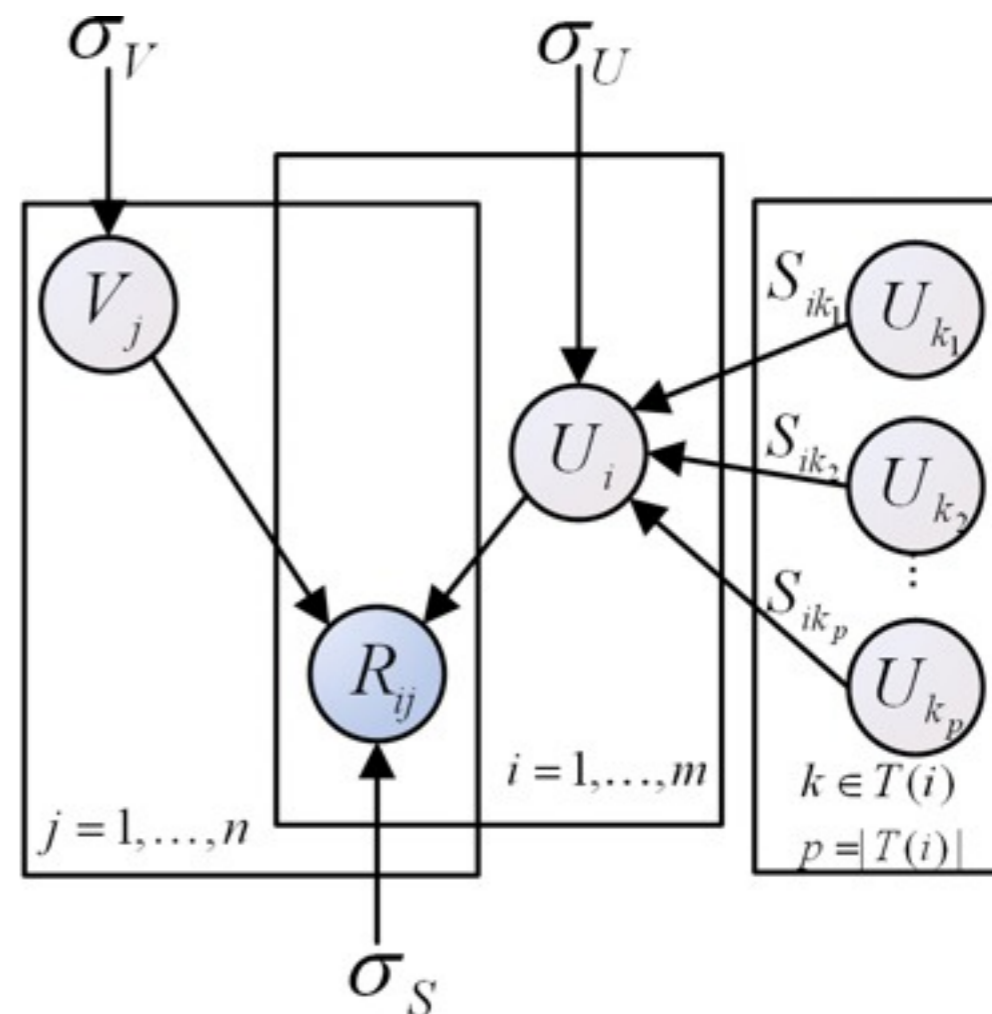
Recommendations by Trusted Friends

$$\hat{R}_{ik} = \frac{\sum_{j \in \mathcal{T}(i)} R_{jk} S_{ij}}{|\mathcal{T}(i)|}$$

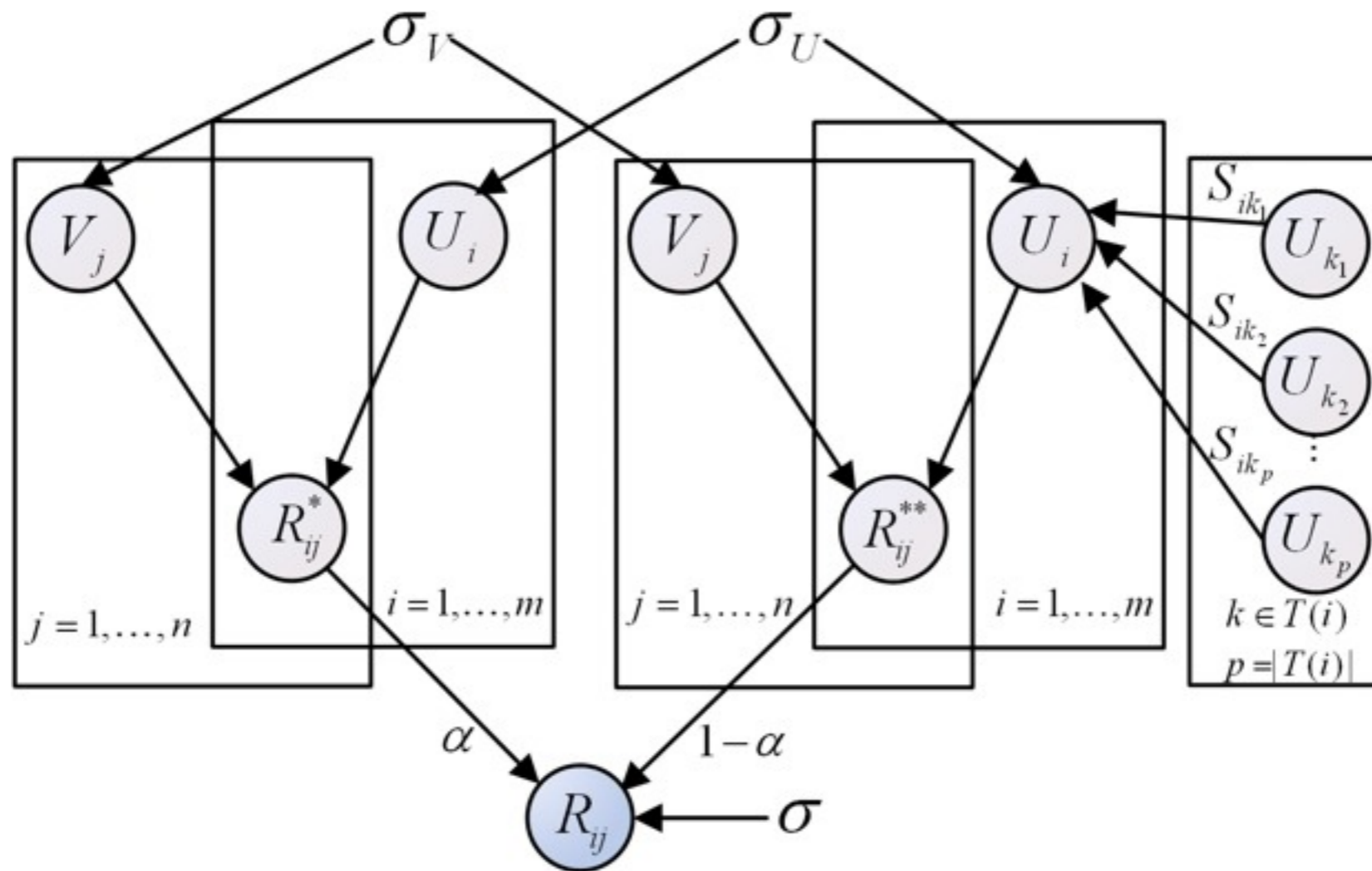
$$\hat{R}_{ik} = \sum_{j \in \mathcal{T}(i)} R_{jk} S_{ij}$$

$$p(R|S, U, V, \sigma_R^2) =$$

$$\prod_{i=1}^m \prod_{j=1}^n \left[\mathcal{N} \left(R_{ij} \mid g \left(\sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j \right), \sigma_S^2 \right) \right]^{I_{ij}^R}$$



Recommendation with Social Trust Ensemble



$$\prod_{i=1}^m \prod_{j=1}^n \left[\mathcal{N} \left(R_{ij} \mid g \left(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T V_j \right), \sigma^2 \right) \right]^{I_{ij}^R}$$



Recommendation with Social Trust Ensemble

$$\begin{aligned}
 \mathcal{L}(R, S, U, V) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j))^2 \\
 &\quad + \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2,
 \end{aligned} \tag{15}$$

$$\begin{aligned}
 \frac{\partial \mathcal{L}}{\partial U_i} &= \alpha \sum_{j=1}^n I_{ij}^R g'(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) V_j \\
 &\quad \times (g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) - R_{ij}) \\
 &\quad + (1 - \alpha) \sum_{p \in \mathcal{B}(i)} \sum_{j=1}^n I_{pj}^R g'(\alpha U_p^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(p)} S_{pk} U_k^T V_j) \\
 &\quad \times (g(\alpha U_p^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(p)} S_{pk} U_k^T V_j) - R_{pj}) S_{pi} V_j \\
 &\quad + \lambda_U U_i,
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial \mathcal{L}}{\partial V_j} &= \sum_{i=1}^m I_{ij}^R g'(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) \\
 &\quad \times (g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) - R_{ij}) \\
 &\quad \times (\alpha U_i + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T) + \lambda_V V_j,
 \end{aligned}$$



Complexity

- In general, the complexity of this method is linear with the observations the user-item matrix



Epinions Dataset

- 51,670 users who rated 83,509 items with totally 631,064 ratings
- Rating Density 0.015%
- The total number of issued trust statements is 511,799



Metrics

- Mean Absolute Error and Root Mean Square Error

$$MAE = \frac{\sum_{i,j} |r_{i,j} - \hat{r}_{i,j}|}{N}$$

$$RMSE = \sqrt{\frac{\sum_{i,j} (r_{i,j} - \hat{r}_{i,j})^2}{N}}$$



Comparisons

Table III: Performance Comparisons (A Smaller MAE or RMSE Value Means a Better Performance)

Training Data	Metrics	Dimensionality = 5						
		UserMean	ItemMean	NMF	PMF	Trust	SoRec	RSTE
90%	MAE	0.9134	0.9768	0.8738	0.8676	0.9054	0.8442	0.8377
	RMSE	1.1688	1.2375	1.1649	1.1575	1.1959	1.1333	1.1109
80%	MAE	0.9285	0.9913	0.8975	0.8951	0.9221	0.8638	0.8594
	RMSE	1.1817	1.2584	1.1861	1.1826	1.2140	1.1530	1.1346

Training Data	Metrics	Dimensionality = 10						
		UserMean	ItemMean	NMF	PMF	Trust	SoRec	RSTE
90%	MAE	0.9134	0.9768	0.8712	0.8651	0.9039	0.8404	0.8367
	RMSE	1.1688	1.2375	1.1621	1.1544	1.1917	1.1293	1.1094
80%	MAE	0.9285	0.9913	0.8951	0.8886	0.9215	0.8580	0.8537
	RMSE	1.1817	1.2584	1.1832	1.1760	1.2132	1.1492	1.1256

NMF --- D. D. Lee and H. S. Seung (Nature 1999)

PMF --- R. Salakhutdinov and A. Mnih (NIPS 2008)

SoRec --- H. Ma, H. Yang, M. R. Lyu and I. King (CIKM 2008)

Trust, RSTE --- H. Ma, I. King and M. R. Lyu (SIGIR 2009)

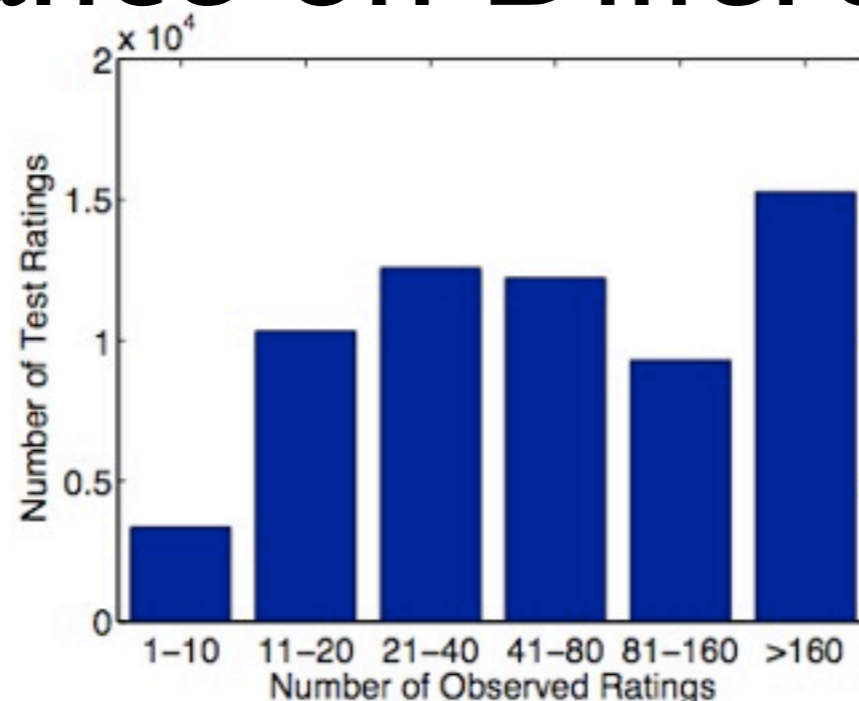


Performance on Different Users

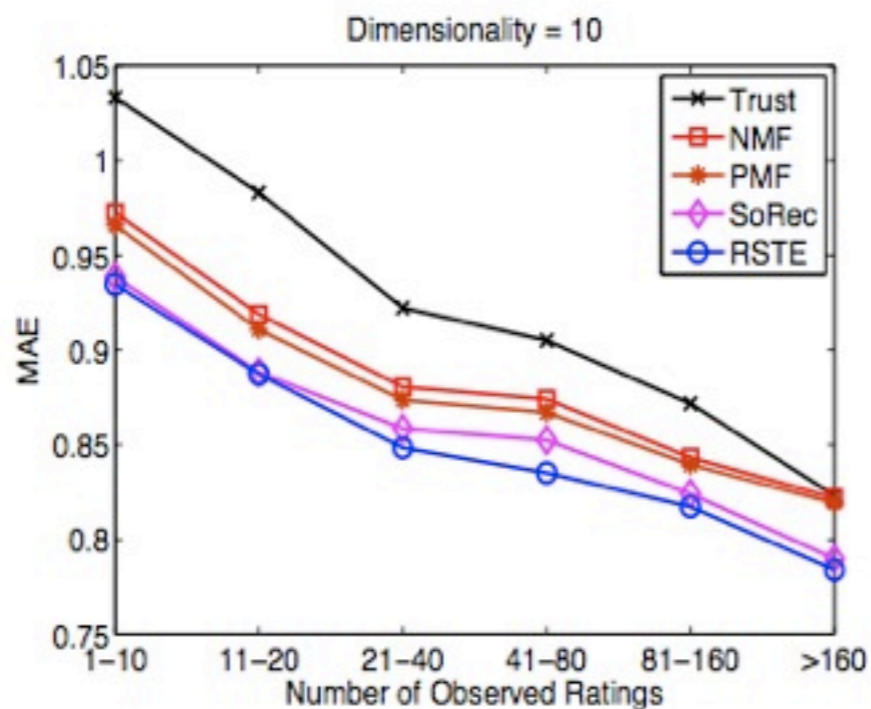
- Group all the users based on the number of observed ratings in the training data
- 6 classes: “1 – 10”, “11 – 20”, “21 – 40”, “41 – 80”, “81 – 160”, “> 160”,



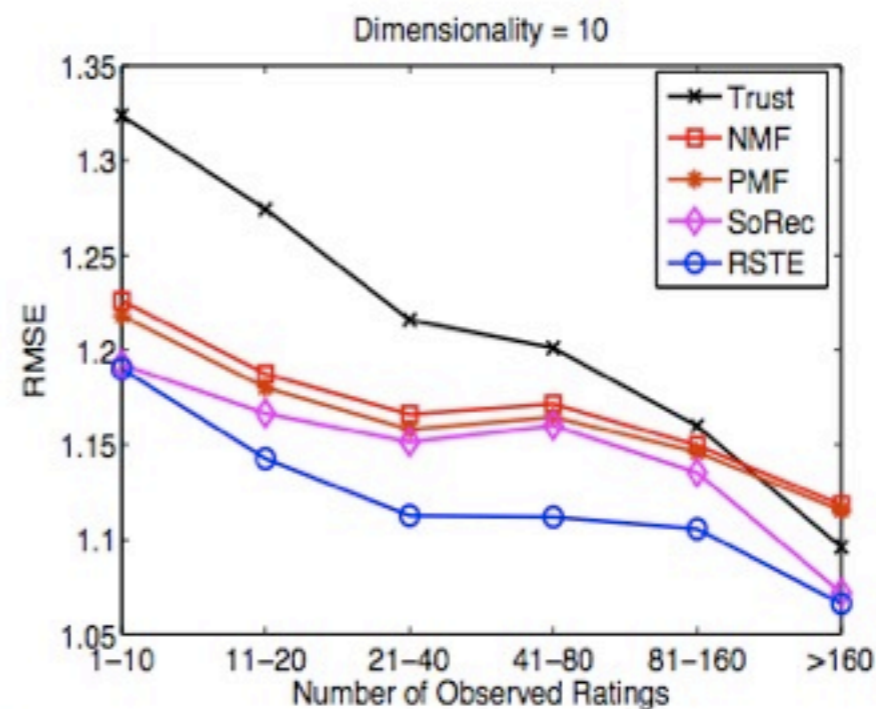
Performance on Different Users



(a) Distribution of Testing Data (90% as Training Data)



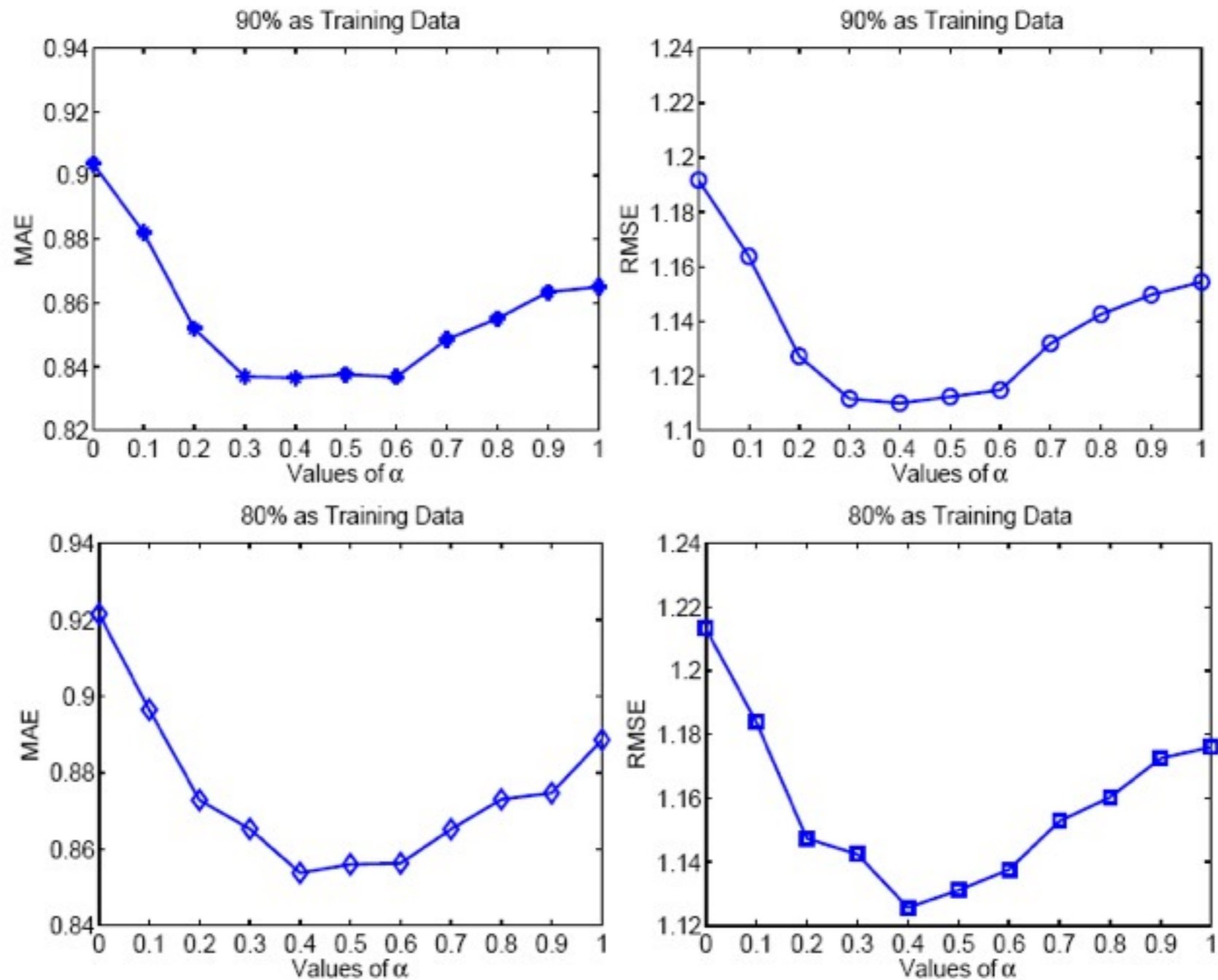
(b) MAE Comparison on Different User Rating Scales (90% as Training Data)



(c) RMSE Comparison on Different User Rating Scales (90% as Training Data)



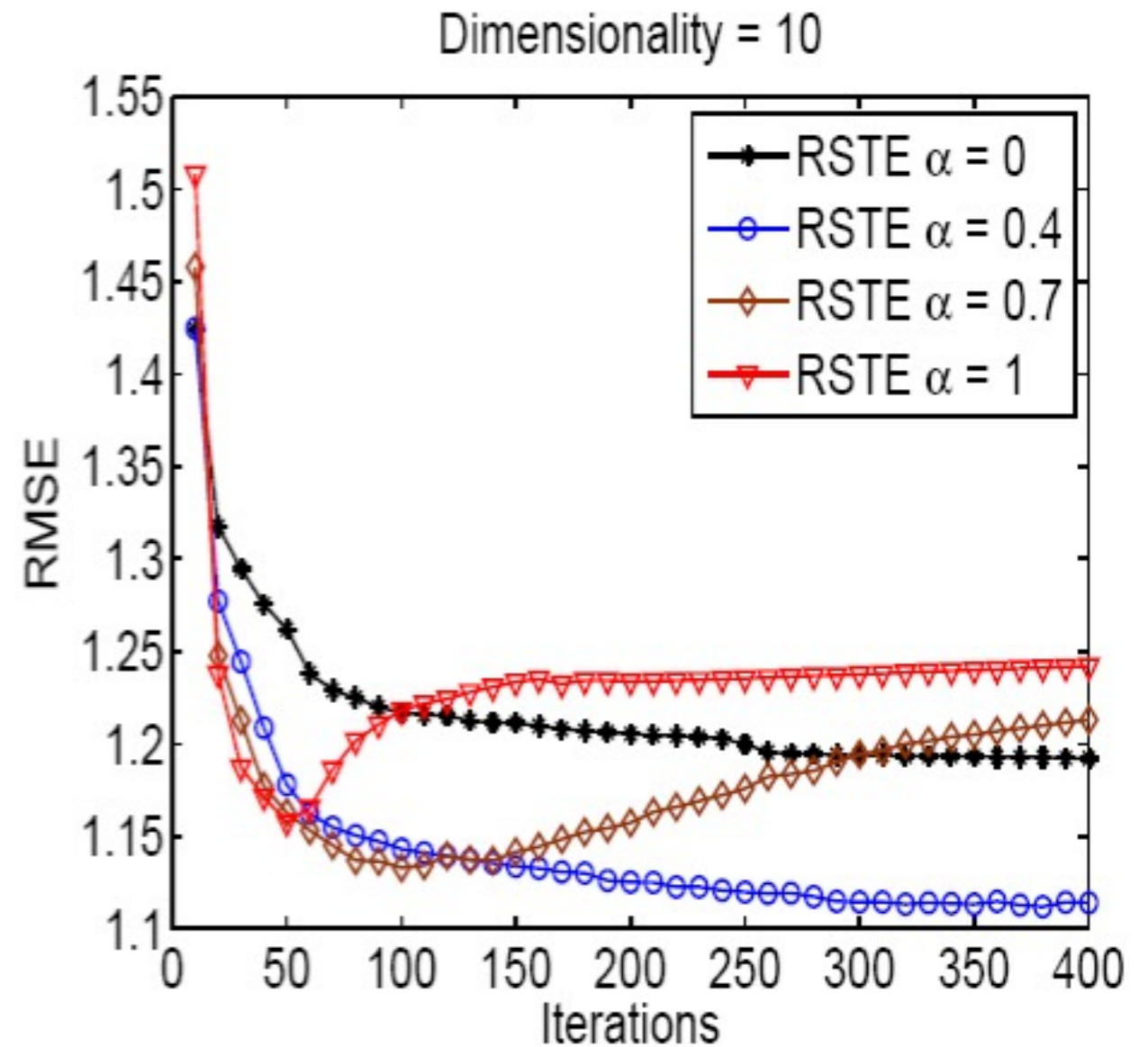
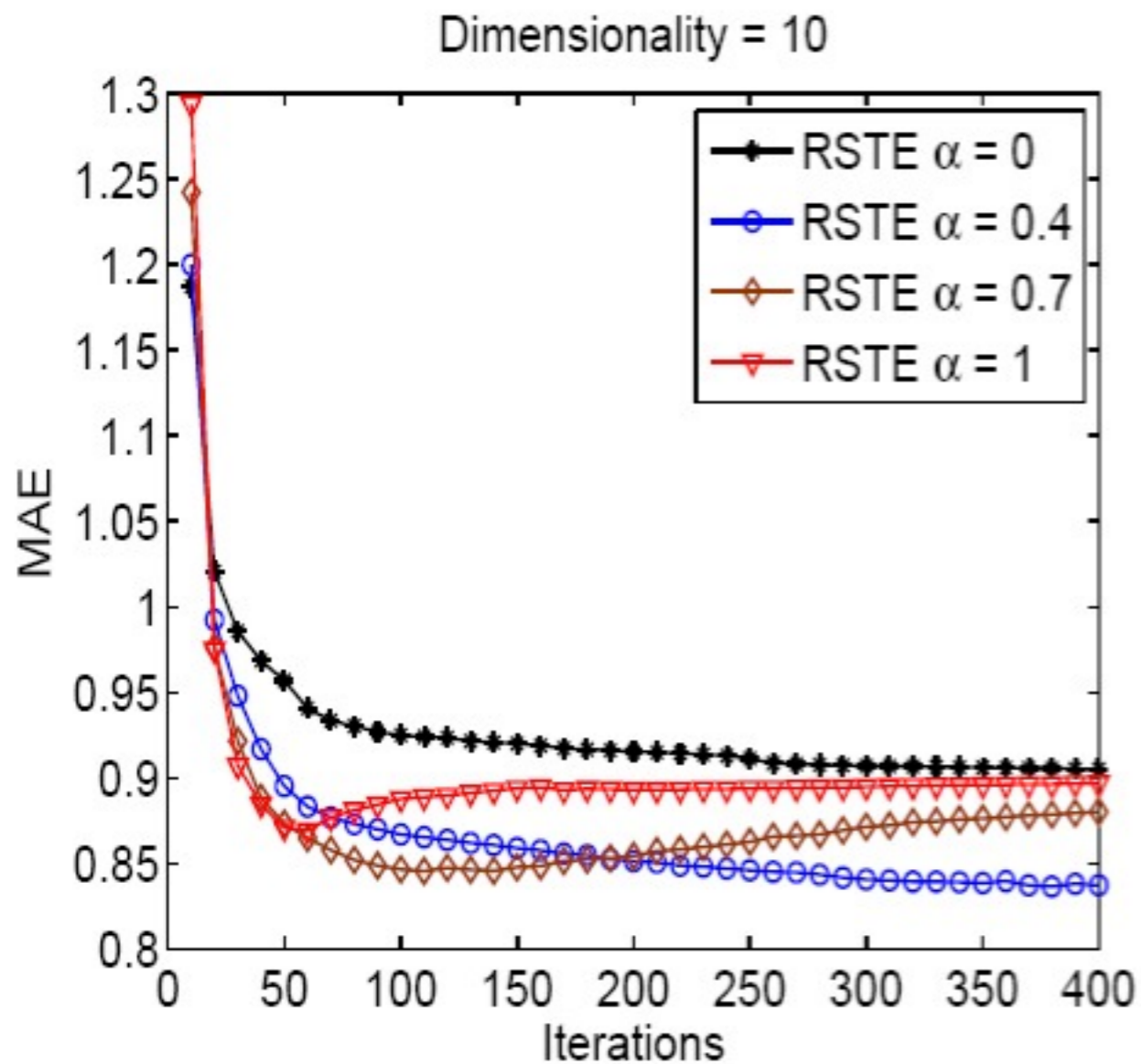
Impact of Parameter Alpha



Impact of Parameter α (Dimensionality = 10)



MAE and RMSE Changes with Iterations

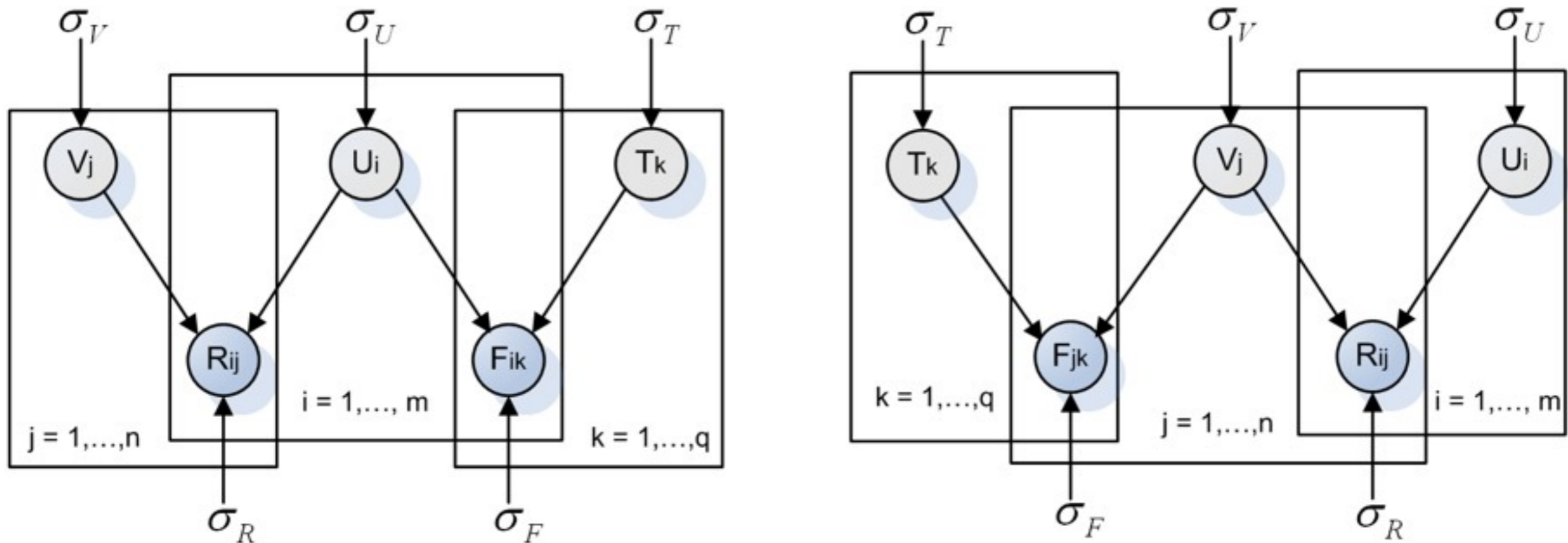


90% as Training Data



Further Discussion of SoRec

- Improving Recommender Systems Using Social Tags



MovieLens Dataset

71,567 users, **10,681** movies,
10,000,054 ratings, **95,580** tags



Further Discussion of SoRec

- MAE

Table V: MAE comparison with other approaches on MovieLens dataset
(A smaller MAE value means a better performance)

Methods		80% Training	50% Training	30% Training	10% Training
User Mean		0.7686	0.7710	0.7742	0.8234
Item Mean		0.7379	0.7389	0.7399	0.7484
5D	SVD	0.6390	0.6547	0.6707	0.7448
	PMF	0.6325	0.6542	0.6698	0.7430
	SoRecUser	0.6209	0.6419	0.6607	0.7040
	SoRecItem	0.6199	0.6407	0.6395	0.7026
10D	SVD	0.6386	0.6534	0.6693	0.7431
	PMF	0.6312	0.6530	0.6683	0.7417
	SoRecUser	0.6197	0.6408	0.6595	0.7028
	SoRecItem	0.6187	0.6395	0.6584	0.7016



Further Discussion of SoRec

- RMSE

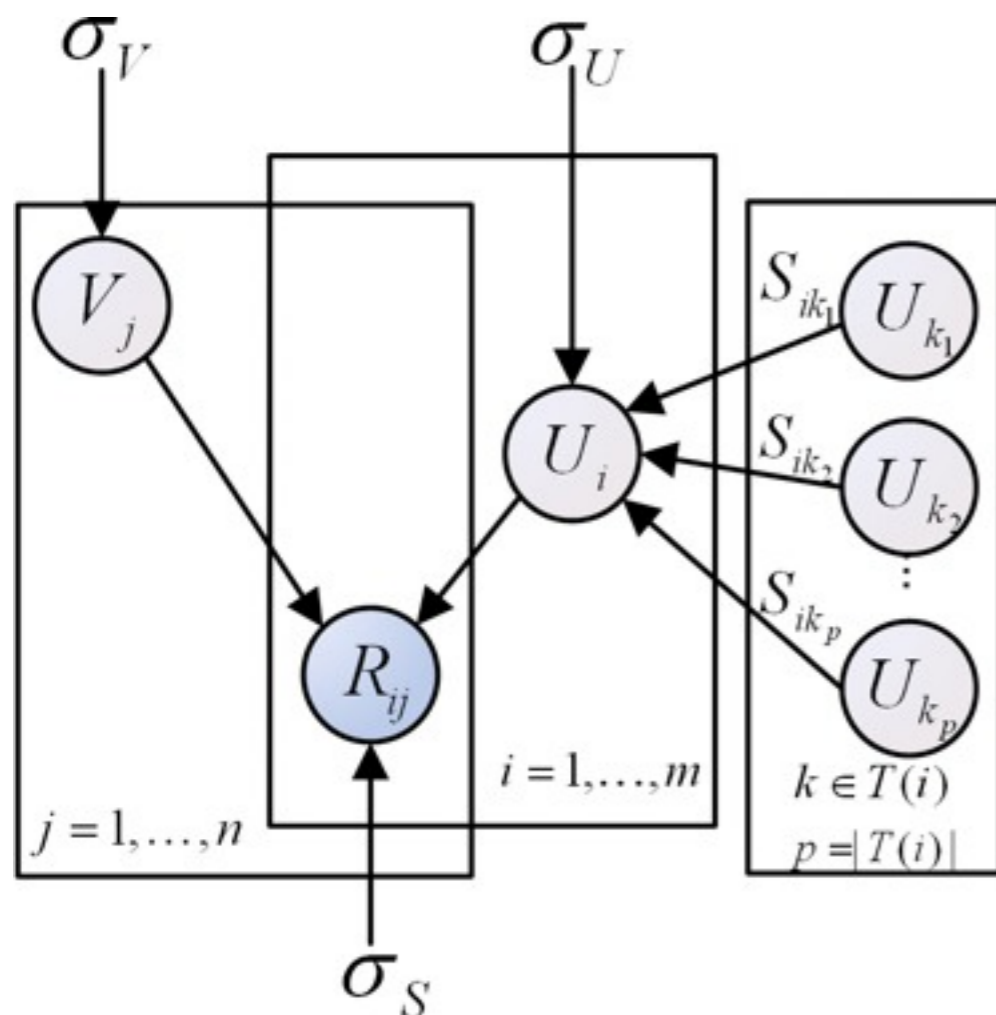
Table VI: RMSE comparison with other approaches on MovieLens dataset (A smaller RMSE value means a better performance)

Methods		80% Training	50% Training	30% Training	10% Training
User Mean		0.9779	0.9816	0.9869	1.1587
Item Mean		0.9440	0.9463	0.9505	0.9851
5D	SVD	0.8327	0.8524	0.8743	0.9892
	PMF	0.8310	0.8582	0.8758	0.9698
	SoRecUser	0.8121	0.8384	0.8604	0.9042
	SoRecItem	0.8112	0.8370	0.8591	0.9033
10D	SVD	0.8312	0.8509	0.8728	0.9878
	PMF	0.8295	0.8569	0.8743	0.9681
	SoRecUser	0.8110	0.8372	0.8593	0.9034
	SoRecItem	0.8097	0.8359	0.8578	0.9019



Further Discussion of RSTE

- Relationship with Neighborhood-based methods



- The trusted friends are actually the explicit neighbors
- We can easily apply this method to include implicit neighbors
- Using PCC to calculate similar users for every user

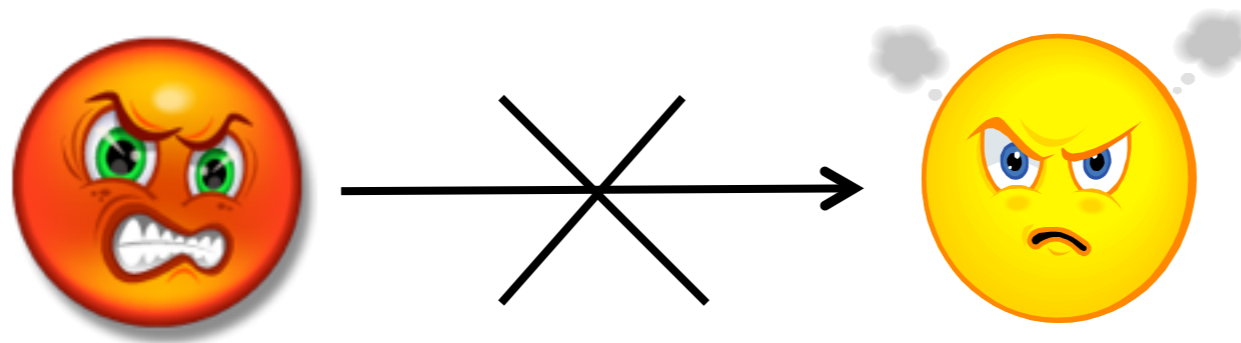


What We Cannot Model Using SoRec and RSTE?

- Propagation of trust



- Distrust



Recommend with Social Distrust

[Hao Ma, et al., RecSys2009]



Distrust

- Users' **distrust** relations can be interpreted as the “**dissimilar**” relations
 - On the web, user U_i distrusts user U_d indicates that user U_i **disagrees** with most of the opinions issued by user U_d .



Distrust

$$\max_U \frac{1}{2} \sum_{i=1}^m \sum_{d \in \mathcal{D}^+(i)} S_{id}^{\mathcal{D}} \|U_i - U_d\|_F^2$$

$$\begin{aligned} \min_{U,V} \mathcal{L}_{\mathcal{D}}(R, S^{\mathcal{D}}, U, V) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(U_i^T V_j))^2 \\ &+ \frac{\beta}{2} \sum_{i=1}^m \sum_{d \in \mathcal{D}^+(i)} (-S_{id}^{\mathcal{D}} \|U_i - U_d\|_F^2) \\ &+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2. \end{aligned}$$



Trust

- Users' **trust** relations can be interpreted as the “**similar**” relations
 - On the web, user U_i trusts user U_t indicates that user U_i **agrees** with most of the opinions issued by user U_t .



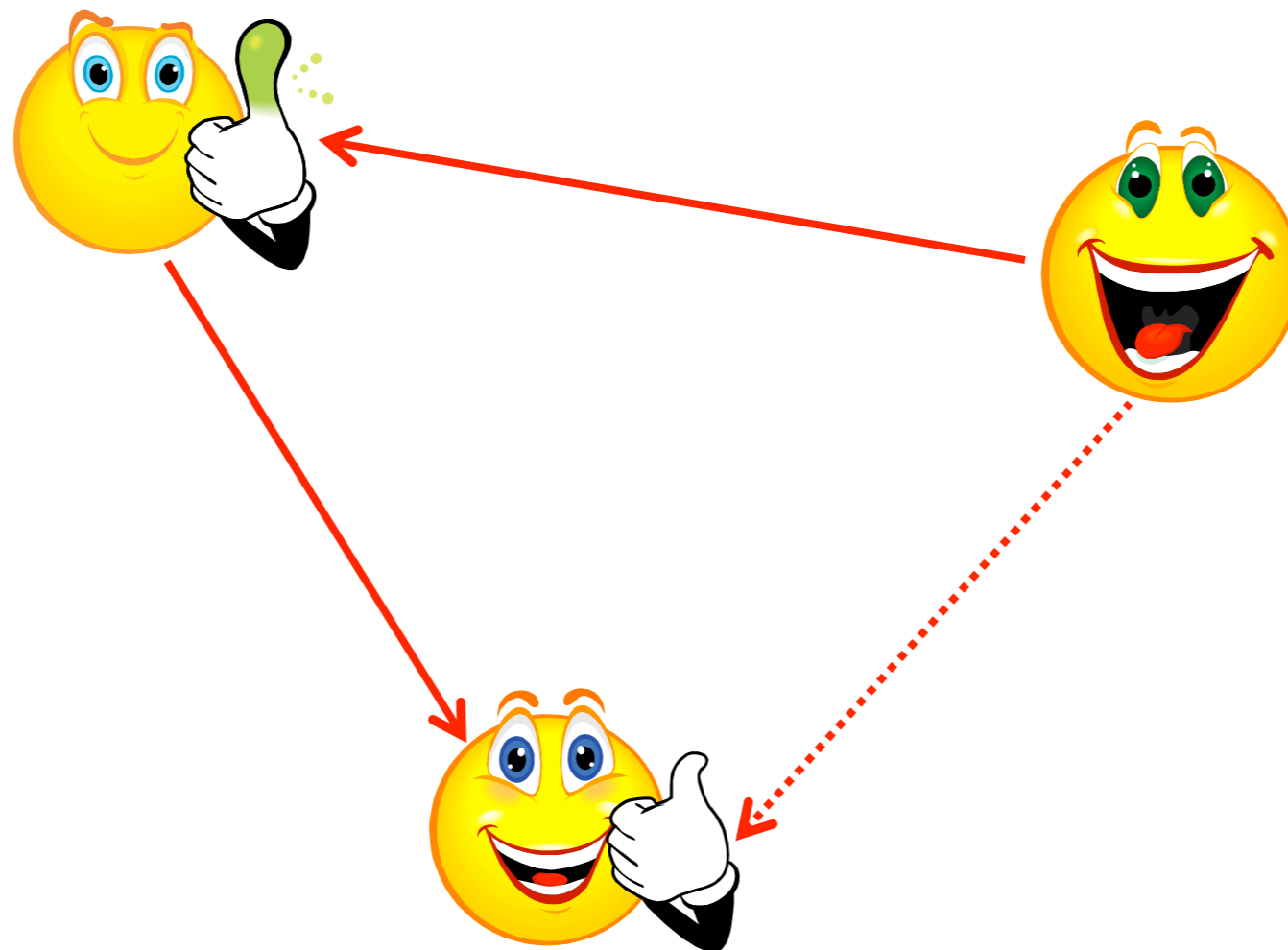
Trust

$$\min_U \frac{1}{2} \sum_{i=1}^m \sum_{t \in \mathcal{T}^+(i)} S_{it}^T \|U_i - U_t\|_F^2$$

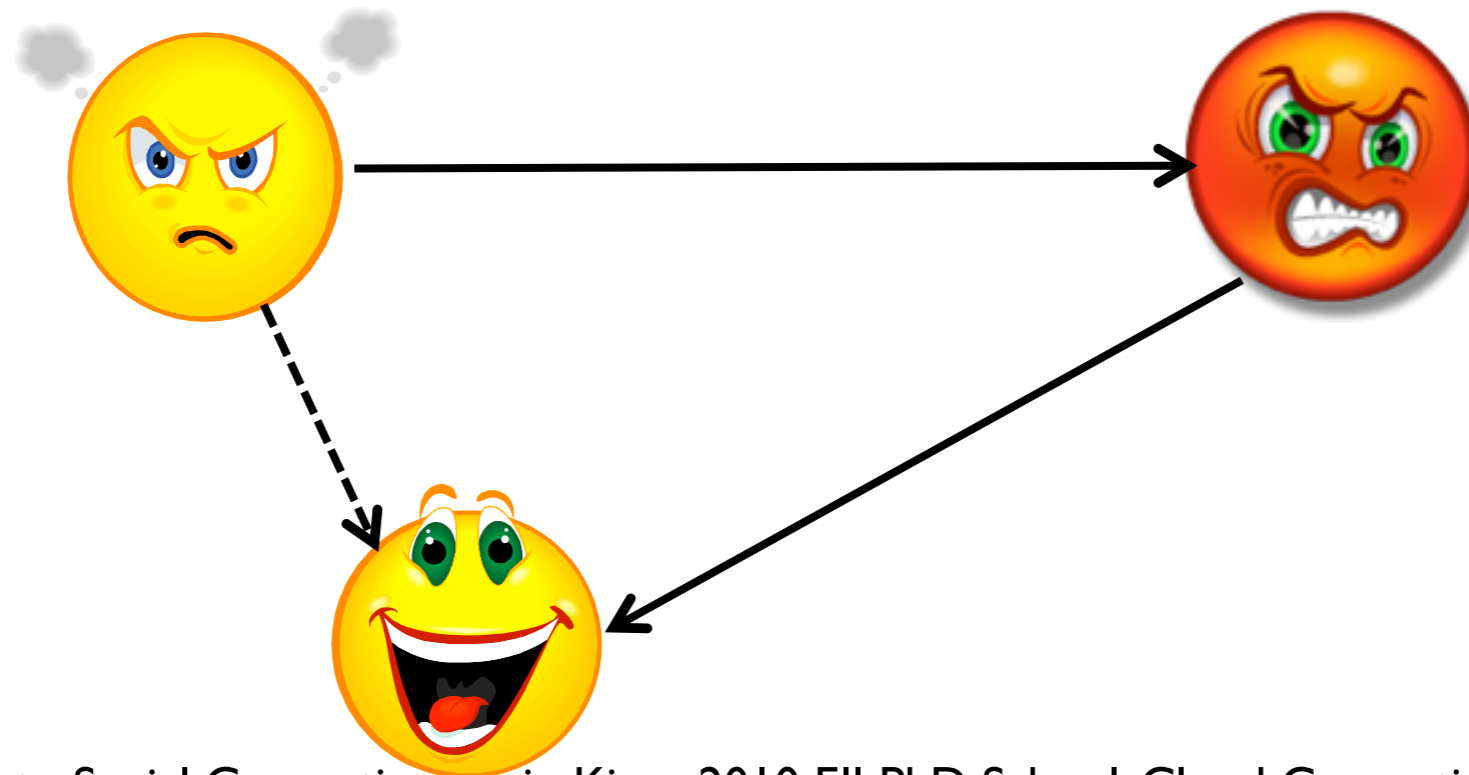
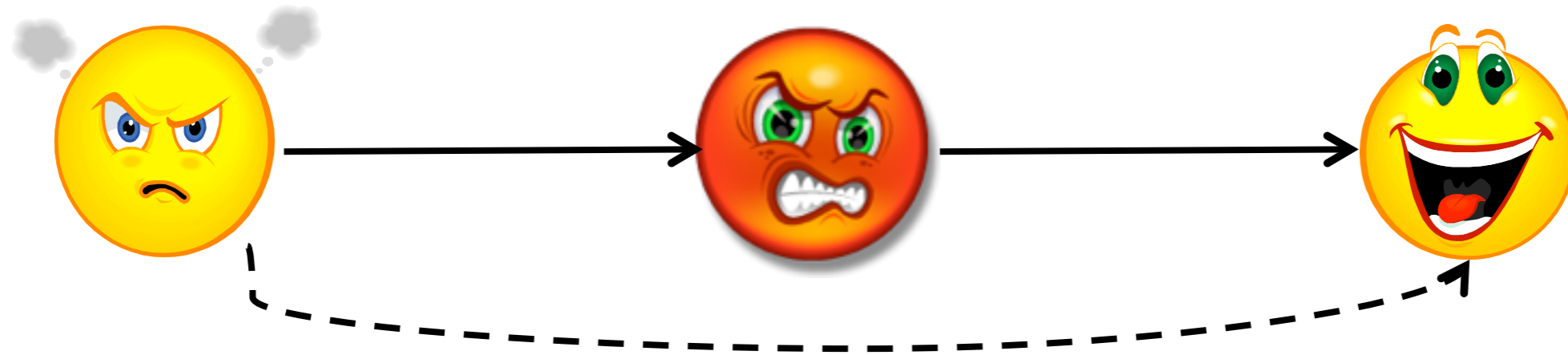
$$\begin{aligned} \min_{U,V} \mathcal{L}_{\mathcal{T}}(R, S^T, U, V) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(U_i^T V_j))^2 \\ &+ \frac{\alpha}{2} \sum_{i=1}^m \sum_{t \in \mathcal{T}^+(i)} (S_{it}^T \|U_i - U_t\|_F^2) \\ &+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2. \end{aligned}$$



Trust Propagation



Distrust Propagation?



Experiments

- Dataset - Epinions
- 131,580 users, 755,137 items, 13,430,209 ratings
- 717,129 trust relations, 123,670 distrust relations



Data Statistics

Table 1: Statistics of User-Item Rating Matrix of Epinions

Statistics	User	Item
Min. Num. of Ratings	1	1
Max. Num. of Ratings	162169	1179
Avg. Num. of Ratings	102.07	17.79

Table 2: Statistics of Trust Network of Epinions

Statistics	Trust per User	Be Trusted per User
Max. Num.	2070	3338
Avg. Num.	5.45	5.45

Table 3: Statistics of Distrust Network of Epinions

Statistics	Distrust per User	Be Distrusted per User
Max. Num.	1562	540
Avg. Num.	0.94	0.94



Experiments

RMSE

Dataset	Traning Data	Dimensionality	PMF	SoRec	RWD	RWT
Epinions	5%	5D	1.228	1.199	1.186	1.177
		10D	1.214	1.198	1.185	1.176
	10%	5D	0.990	0.944	0.932	0.924
		10D	0.977	0.941	0.931	0.923
	20%	5D	0.819	0.788	0.723	0.721
		10D	0.818	0.787	0.723	0.720



Impact of Parameters

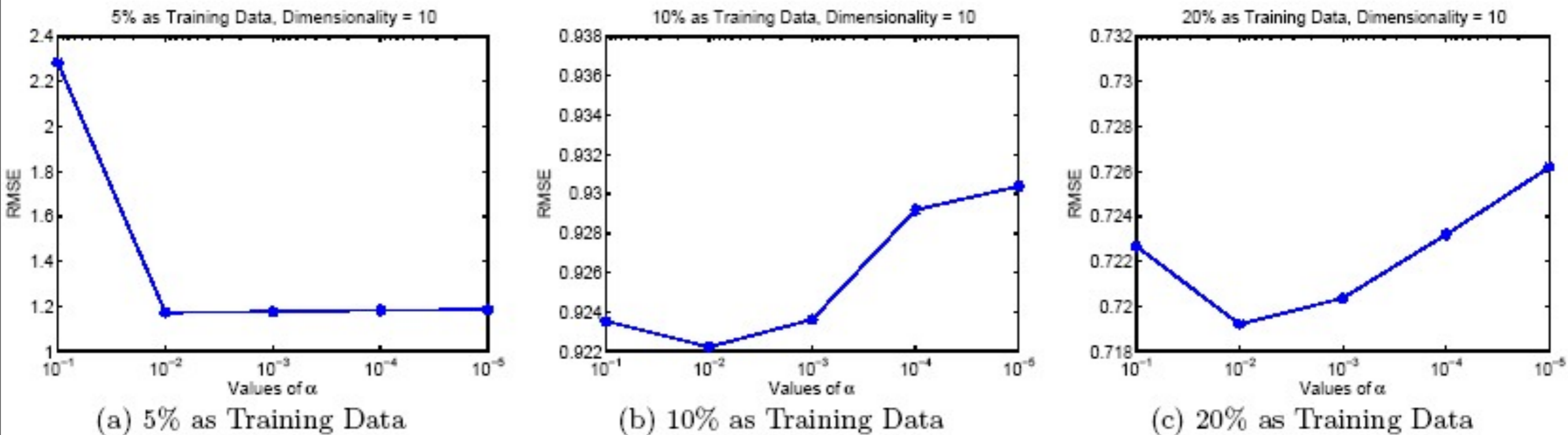


Figure 6: Impact of Parameter α

Alpha = 0.01 will get the best performance!
Parameter beta basically shares the same trend!



Social Recommender Systems

- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- **Social-based Recommender Systems**



Comparison

- Trust-aware Recommender systems

- Trust network
- Trust relations can be treated as “similar” relations
- Few dataset available on the web

- Social-based Recommender Systems

- Social friend network, mutual relations
- Friends are very divers, and may have different tastes
- Lots of web sites have social network implementation



References

- J. Basilico and T. Hofmann. Unifying collaborative and content-based filtering. In ICML, 2004.
- J. S. Breese, D. Heckerman, and C. M. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In UAI, pages 43–52, 1998.
- M. Deshpande and G. Karypis. Item-based top-N recommendation algorithms. ACM Trans. Inf. Syst., 22(1):143–177, 2004.
- J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl. An algorithmic framework for performing collaborative filtering. In SIGIR, pages 230–237. ACM, 1999.
- J. L. Herlocker, J. A. Konstan, and J. Riedl. An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. Inf. Retr., 5(4):287–310, 2002.
- G. Linden, B. Smith, and J. York. Industry report: Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Distributed Systems Online, 4(1), 2003.



References

- H. Ma, I. King, and M. R. Lyu. Effective missing data prediction for collaborative filtering. In SIGIR, pages 39–46, 2007.
- H. Ma, H. Yang, M. R. Lyu, and I. King. SoRec: social recommendation using probabilistic matrix factorization. In CIKM, pages 931–940, 2008.
- H. Ma, I. King, and M. R. Lyu. Learning to recommend with social trust ensemble. In SIGIR, pages 203–210, 2009.
- H. Ma, M. R. Lyu, and I. King. Learning to recommend with trust and distrust relationships. In RecSys, pages 189–196, 2009.
- B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In WWW, pages 285–295, 2001.
- R. Salakhutdinov, and A. Mnih. Probabilistic Matrix Factorization. In NIPS, 2007.
- D. D. Lee, and H. S. Seung. Algorithms for Non-negative Matrix Factorization. In NIPS, pages 556–562, 2000.



Outline

- Introduction
- Social Search Engine
- Social Recommender Systems
- **Social Media Analysis**



Social Media Analysis

- Social Media Ranking
- Tag Recommendation
- News Recommendation
- User Recommendation
- Twitter-powered Recommendation



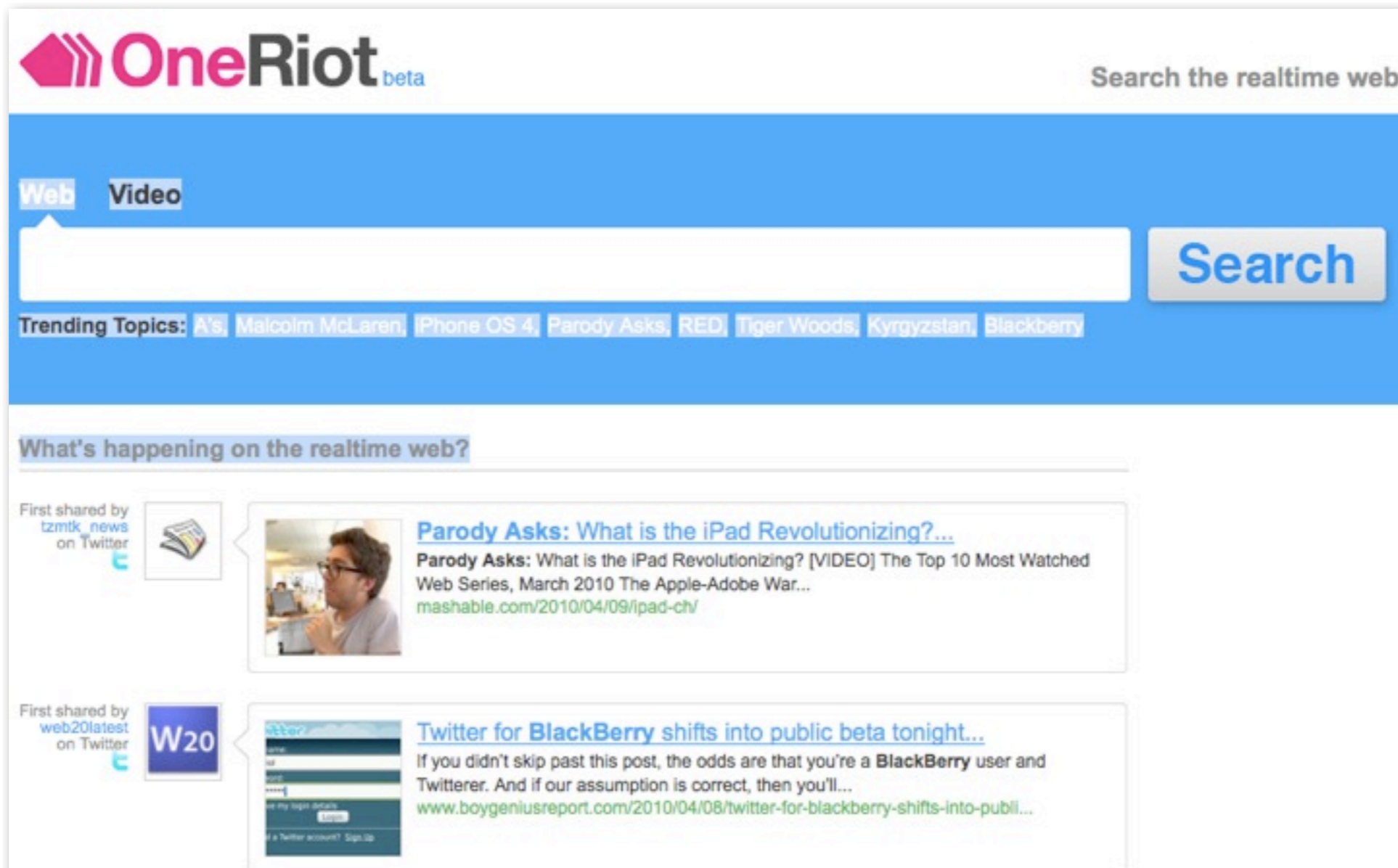
Social Media Ranking

- Pulse Rank - OneRiot
- Reddit Algorithm
- Digg Algorithm
- Google's Page Rank



Pulse Rank - OneRiot

- A realtime web search engine, which archives and makes searchable news, videos and blogs being discussed on the web, ordered to reflect current social relevance.



The screenshot shows the OneRiot website interface. At the top left is the OneRiot logo with a pink house icon and the word "beta" in blue. To the right is the text "Search the realtime web". Below this is a blue navigation bar with "Web" and "Video" tabs. A large white search input field is centered, with a blue "Search" button to its right. Below the search bar is a "Trending Topics" section with tags for "A's", "Malcolm McLaren", "iPhone OS 4", "Parody Asks", "RED", "Tiger Woods", "Kyrgyzstan", and "Blackberry". The main content area is titled "What's happening on the realtime web?". It features two article cards. The first card is titled "Parody Asks: What is the iPad Revolutionizing?..." and includes a video thumbnail of a man speaking. The second card is titled "Twitter for BlackBerry shifts into public beta tonight..." and includes a thumbnail of the Twitter login page.



Pulse Rank - OneRiot

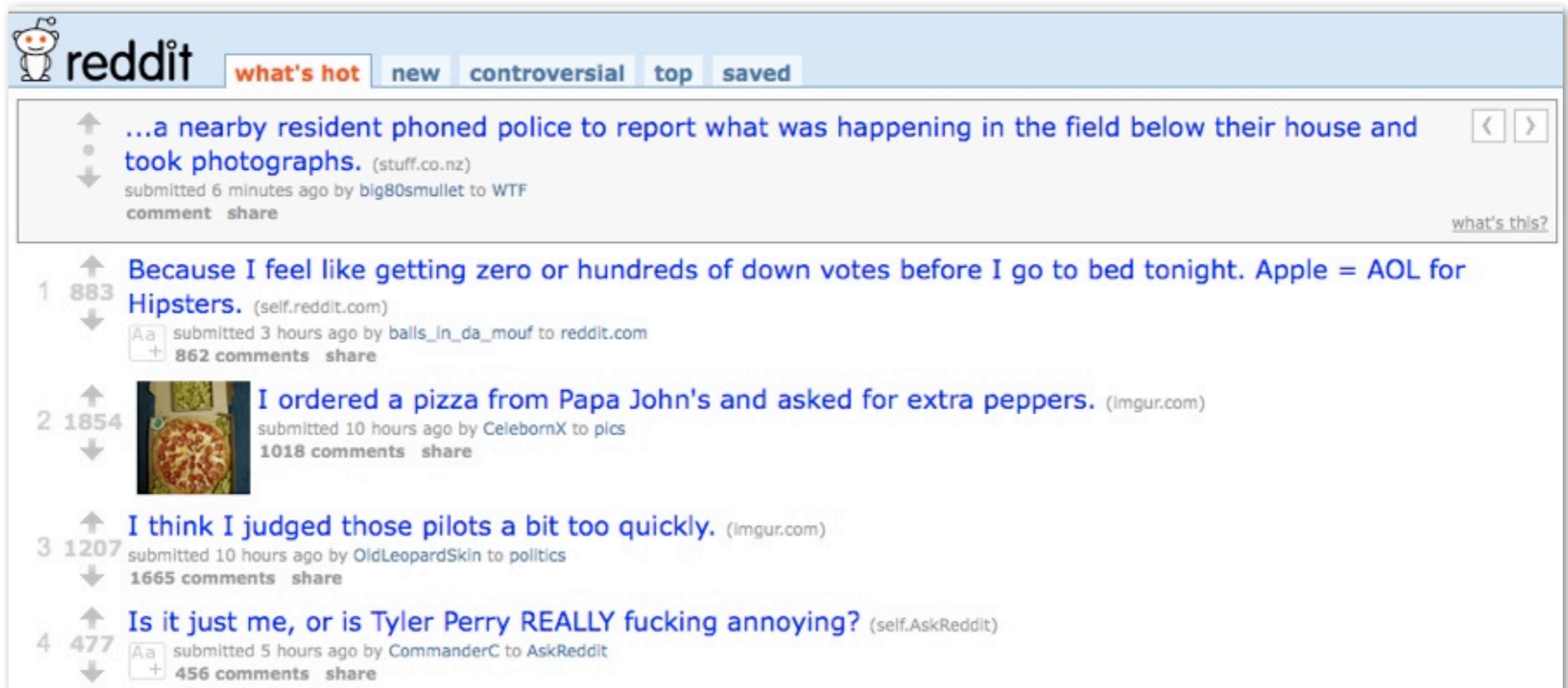
- “Pulse Rank” algorithm looks at dozens of factors that give “weight” to certain results
- **Freshness:** Is the most recently published content necessarily the most relevant?
- **Domain Authority:** An article about Obama on New York Times should weight higher than the article on my blog.
- **People Authority:** Who is sharing this link on the social web?
- **Acceleration:** Is this page increasing in hotness or decreasing in hotness?

From <http://blog.oneriot.com/content/2009/06/oneriot-pulse-rank/>



Reddit Algorithm

- **Reddit** is a social news website on which users can post links to content on the Internet. Other users may then vote the posted links up or down, causing them to become more or less prominent on the reddit home page.



The screenshot shows the Reddit homepage with the 'what's hot' tab selected. The top navigation bar includes the Reddit logo and tabs for 'what's hot', 'new', 'controversial', 'top', and 'saved'. The main content area displays a list of posts:

- Post 1: "...a nearby resident phoned police to report what was happening in the field below their house and took photographs." (stuff.co.nz) submitted 6 minutes ago by big80smullet to WTF. 883 upvotes, 862 comments.
- Post 2: "Because I feel like getting zero or hundreds of down votes before I go to bed tonight. Apple = AOL for Hipsters." (self.reddit.com) submitted 3 hours ago by balls_in_da_mouf to reddit.com. 1854 upvotes, 1018 comments.
- Post 3: "I ordered a pizza from Papa John's and asked for extra peppers." (imgur.com) submitted 10 hours ago by CelebornX to pics. 1207 upvotes, 1665 comments.
- Post 4: "I think I judged those pilots a bit too quickly." (imgur.com) submitted 10 hours ago by OldLeopardSkin to politics. 477 upvotes, 456 comments.
- Post 5: "Is it just me, or is Tyler Perry REALLY fucking annoying?" (self.AskReddit) submitted 5 hours ago by CommanderC to AskReddit. 456 comments.



Reddit Algorithm

- Time differences

$$t_s = A - B$$

- Differences of the up votes and down votes

$$x = U - D$$

$$y = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \quad z = \begin{cases} |x| & \text{if } |x| \geq 1 \\ 1 & \text{if } |x| < 1 \end{cases}$$

- Ranking functions

$$f(t_s, y, z) = \log_{10} z + \frac{yt_s}{45000}$$

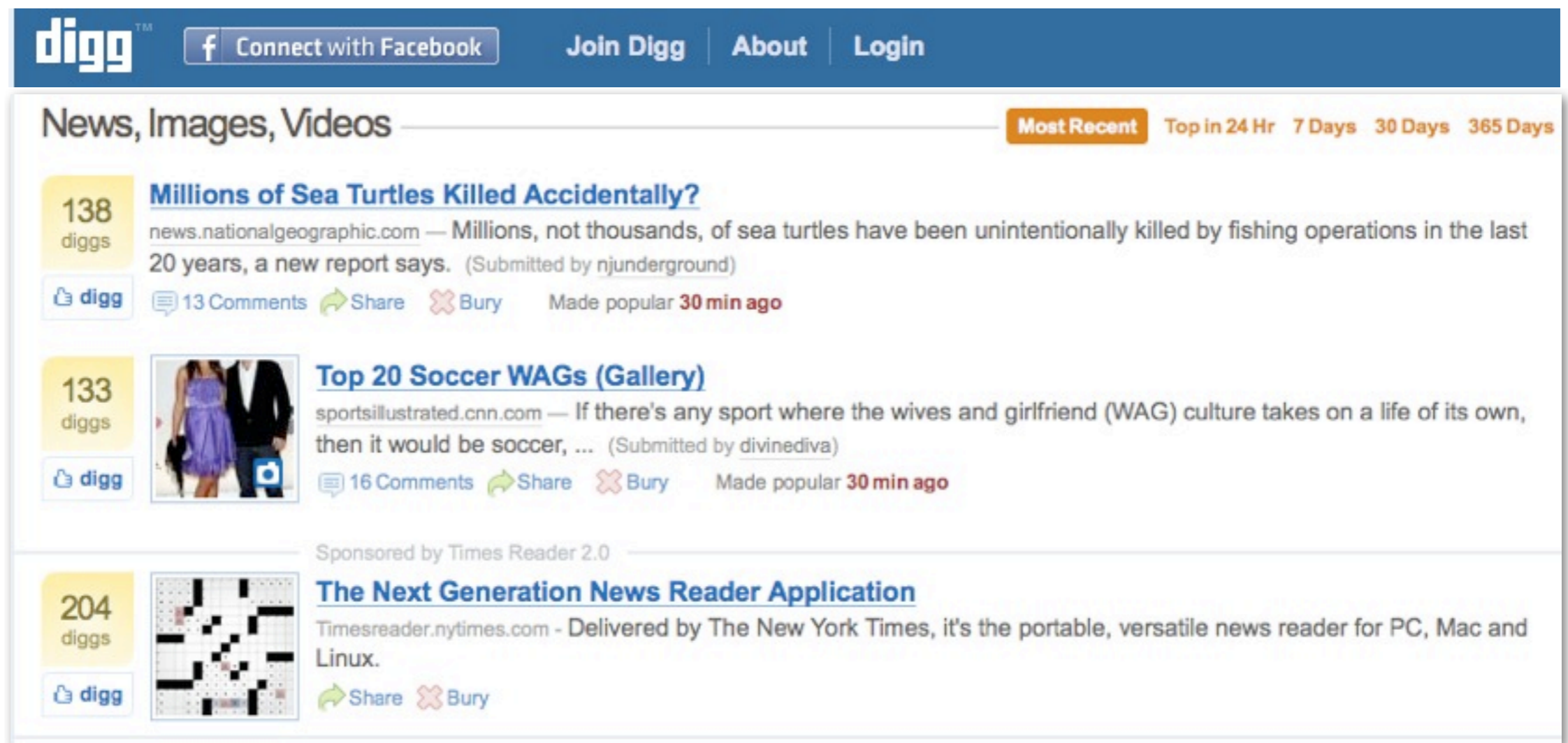
From <http://uggedal.com/reddit.cf.algorithm.png>

Introduction to Social Computing, Irwin King, 2010 ELL PhD School: Cloud Computing, Service Computing & Social Networks, November 23-27, 2010, Brisbane, Australia



Digg Algorithm

- A social news website made for people to discover and share content from anywhere on the Internet, by submitting links and stories, and voting and commenting on submitted links and stories



The screenshot shows the Digg website interface. At the top, there is a blue navigation bar with the Digg logo, a Facebook connection button, and links for 'Join Digg', 'About', and 'Login'. Below the navigation bar, the main content area is titled 'News, Images, Videos' and includes sorting options: 'Most Recent', 'Top in 24 Hr', '7 Days', '30 Days', and '365 Days'. Three news items are displayed:

- 138 diggs**: [Millions of Sea Turtles Killed Accidentally?](#) from news.nationalgeographic.com. Description: 'Millions, not thousands, of sea turtles have been unintentionally killed by fishing operations in the last 20 years, a new report says.' (Submitted by njunderground). Interaction: 13 Comments, Share, Bury, Made popular 30 min ago.
- 133 diggs**: [Top 20 Soccer WAGs \(Gallery\)](#) from sportsillustrated.cnn.com. Description: 'If there's any sport where the wives and girlfriend (WAG) culture takes on a life of its own, then it would be soccer, ...' (Submitted by divinediva). Interaction: 16 Comments, Share, Bury, Made popular 30 min ago.
- 204 diggs**: [The Next Generation News Reader Application](#) from Timesreader.nytimes.com. Description: 'Delivered by The New York Times, it's the portable, versatile news reader for PC, Mac and Linux.' Interaction: Share, Bury.

A sponsored section for 'Times Reader 2.0' is also visible.



Digg Algorithm

- **The rapidity of the votes**
If you get 40-50 votes (no matter what users digg) in the first 30 minutes, you're probably on the frontpage.
- **The rank of the users that vote the article**
The highest it is on the top list, the better.
- **The number of comments, and the positive diggs that each article receives**
If you have a lot of negative rated comments that can hurt more than help actually.
- **The number of buries your story gets**
- **The submitted / promoted stories ratio of the users that vote**
If 12-14 users with at least a 70% ratio, vote your article, you can make the frontpage much easier.



How Google Ranks Tweets

[Latest results for **jesus**](#) - [Pause](#)

Jer: It's gonna be 79 today!? Matt: **Jesus**?

[happyinc77](#) - [Twitter](#) - seconds ago

RT [@alaintha](#): [@kirstiealley](#) happy **jesus** resurection day

[tinytott67](#) - [Twitter](#) - seconds ago

Jesus Christ Noel, dial down the mental would you? It's Deal or No Deal, not Twin Peaks

[doubleshiny](#) - [Twitter](#) - seconds ago

[Latest results for **iphone os4**](#) - [Pause](#)

iPhone OS 4 Event: By The Numbers

[Distimo Blog – iPhone OS 4 Event: By The Numbers](#) - distimo.com

[distimo](#) - [Twitter](#) - 2 minutes ago

Finally awake. Seems like **iPhone OS4** has gripped the world. Oh, and Justin Whats-his-face is still a trending topic.

[jam_ie](#) - [Twitter](#) - 4 minutes ago

[iChat video with front facing camera evidence mounts in iPhone OS ...](#) ☆

0 Apr 2010 - Facebook has announced the iPhone OS 4 SDK developer preview for



How Google Ranks Tweets

- The key is to identify “reputed followers”
- You earn reputation, and then you give reputation
- One user following another in social media is analogous to one page linking to another on the Web. Both are a form of recommendation
- Page Rank on follow graph



Social Media Analysis

- Social Media Ranking
- Tag Recommendation
- News Recommendation
- User Recommendation
- Twitter-powered Recommendation



Why Users Tag?

- Tagging means something specific to the user
- It is easy -- anyone can do it
- Finding things on the Internet
- Serendipitous discovery
- It is social
- New ways to share and discover



Why need Tag Recommendation?

- User tags contain noises
- Automating the tagging process
- Assisting users to tag



Flickr Tag Recommendation based on Collective Knowledge

[B. Sigurbjörnsson, et al., WWW2008]

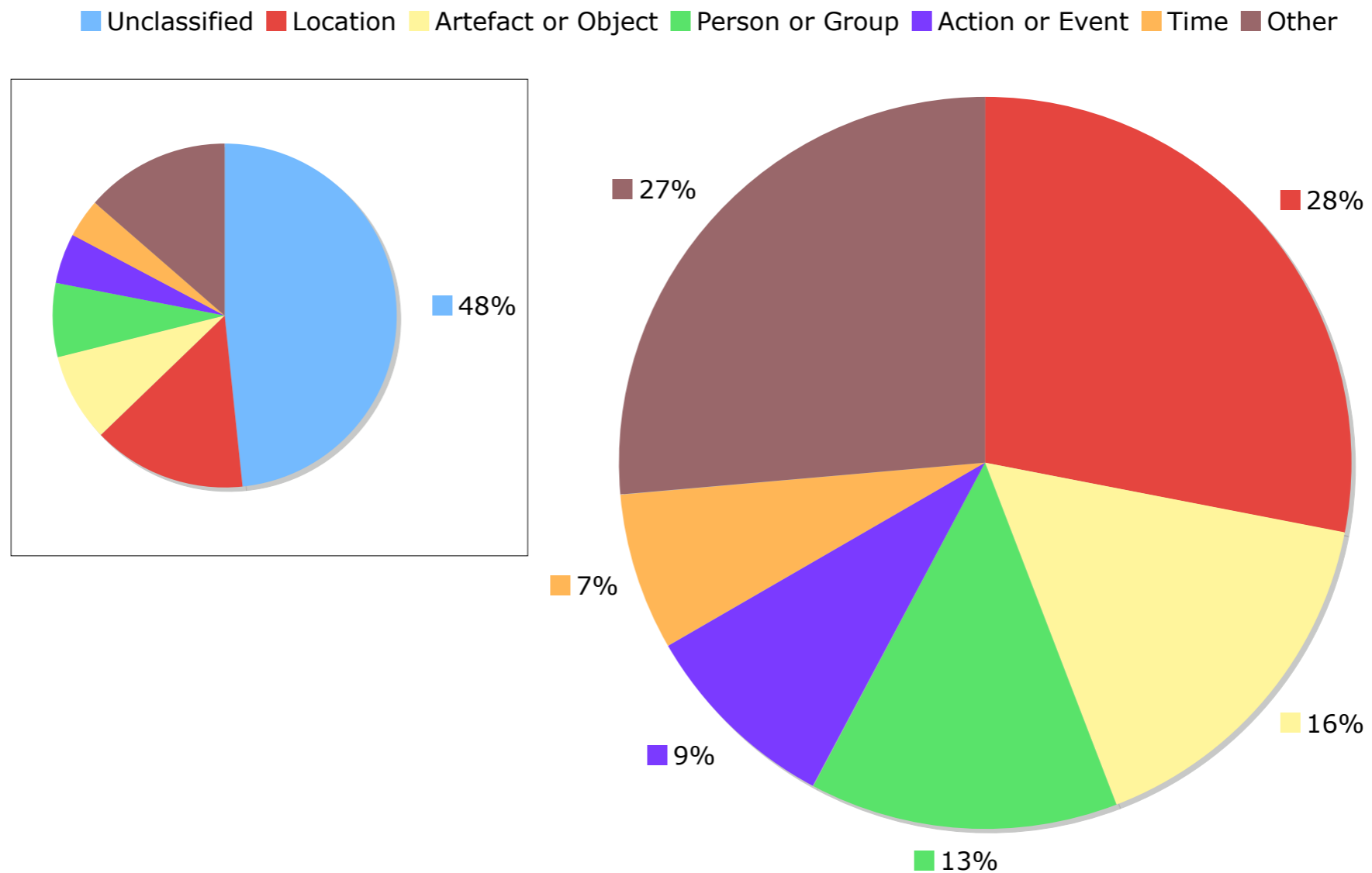


Figure 3: Most frequent WordNet categories for Flickr tags.



Flickr Tag Recommendation based on Collective Knowledge

[B. Sigurbjörnsson, et al., WWW2008]

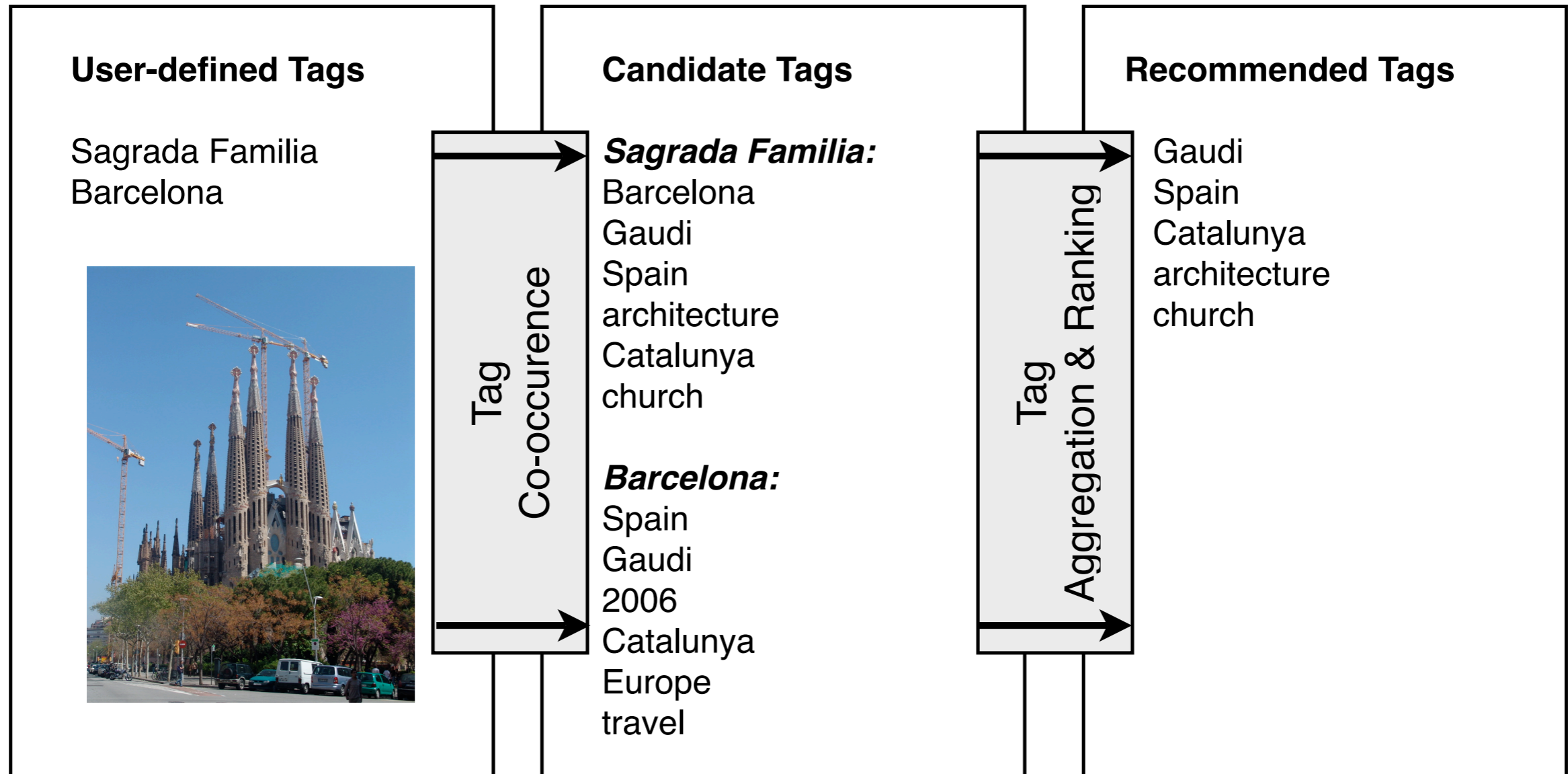
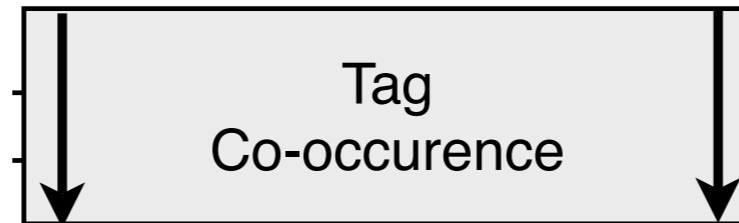


Figure 4: System overview of the tag recommendation process.



Flickr Tag Recommendation based on Collective Knowledge

[B. Sigurbjörnsson, et al., WWW2008]



- Define the Tag Co-occurrence between two tags to be the number of photos where both tags are used in the same annotation
- Symmetric measure: Jaccard Coefficient

$$J(t_i, t_j) := \frac{|t_i \cap t_j|}{|t_i \cup t_j|}$$

- Asymmetric measure:

$$P(t_j | t_i) := \frac{|t_i \cap t_j|}{|t_i|}$$



Flickr Tag Recommendation based on Collective Knowledge

[B. Sigurbjörnsson, et al., WWW2008]

Tag: **Eiffel Tower**



Symmetric Measure:

Tour Eiffel

Eiffel

Seine

La Tour Eiffel

Paris

Good at identifying equivalent tags

Aymmetric Measure:

Paris

France

Tour Eiffel

Eiffel

Europe

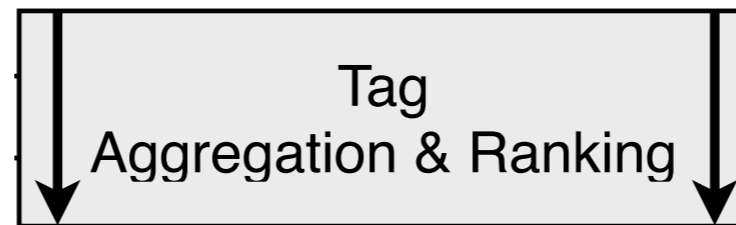
Good at suggesting diverse tags



Flickr Tag Recommendation based on Collective Knowledge

[B. Sigurbjörnsson, et al., WWW2008]

- Aggregation



- Vote

- The voting strategy computes a score for each candidate tag c

$$vote(u, c) = \begin{cases} 1 & \text{if } c \in C_u \\ 0 & \text{otherwise} \end{cases}$$

A score is therefore computed as

$$score(c) := \sum_{u \in U} vote(u, c)$$

- Sum

- The summing strategy sums over the co-occurrence values of the tags

$$score(c) := \sum_{u \in U} (P(c|u)) \quad , \text{if } c \in C_u$$

where $P(c|u)$ calculates the asymmetric co-occurrence values, and u is the user defined tags



Social Media Analysis


- Social Media Ranking
- Tag Recommendation
- **News Recommendation**
- User Recommendation
- Twitter-powered Recommendation



Google News Recommendation


Top Stories

[Silence held across Poland for deceased president](#) ☆
ABC Online - 2 hours ago
Solemnly standing to attention as sirens wailed, Poles fell silent across the country Sunday as they mourned President Lech Kaczynski and top officials killed in a fiery air crash in Russia.
+ Video: [Bells and sirens sound in memory of Polish plane crash victims](#) YouTube RT
[Polish president's body flown home](#) Aljazeera.net
[BBC News](#) - [Xinhua](#) - [The Guardian](#) - [Jewish Telegraphic Agency](#) - [Wikipedia: Lech Kaczyński](#)
[all 5,904 news articles »](#) [Email this story](#)



The Guardian


[Hundreds wounded, 20 killed in Thailand protests](#) ☆
ABC Online - [Mark Willacy](#) - 2 hours ago
The Thai government denies that soldiers fired live bullets into crowds of protesters. (Reuters : Sukree Sukplang) At least 20 people are dead and more than 800 are wounded in Thailand after violent clashes between opposition ...
+ Video: [Thai political crisis turns deadly](#) YouTube Aljazeera.net
[Political Standoff in Bangkok Intensifies](#) New York Times
[Times Online](#) - [Reuters](#) - [The Associated Press](#) - [Wikipedia: National United Front of Democracy](#)
[all 2,174 news articles »](#) [Email this story](#)



Google news


Recommended »

[Pink Preview: Microsoft's Mystery Event](#) ☆
PC World - [Paul Suarez](#) - Apr 10, 2010
Artwork: Chip Taylor Earlier this week Microsoft sent out invitations for a "mystery event" that will take place in San Francisco on Monday.
[Will iPhone 4.0 derail Microsoft's phone plans?](#) CNET
[How iPhone OS destroys Windows Phone 7 without even shipping](#) Ars Technica
[ABC News](#) - [TopNews United States](#) - [Onion Kid](#) - [Fone Arena \(blog\)](#)
[all 83 news articles »](#) [Email this story](#)



PC World

[Staycation Specials: Zip line for free in San Francisco](#) ☆
San Jose Mercury News - [Ann Tatko-Peterson](#) - 8 hours ago
Ride on an urban zip line for free during the British Columbia Experience in San Francisco. At Embarcadero Square, Ziptrek Ecotours has set up a 600-foot zip line that is similar to the popular urban zip line offered to tourists ...
[Reliving the highs of the Vancouver games](#) CNET
[Zip line offers bird's-eye view of city](#) UPI.com



TopNews New Zealand



News Recommendation

- Online news reading has become very popular
- Web provides access to news articles from millions of sources around the world
- Key challenge: help users find the articles that are interesting to read



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

- News click logs analysis
 - Data
 - Google News, over 12-month period, from 2007/07/01 to 2008/06/30
 - Randomly sampled 16,848 users from users who made at least 10 clicks per month
 - Users are from more than 10 different countries and regions



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

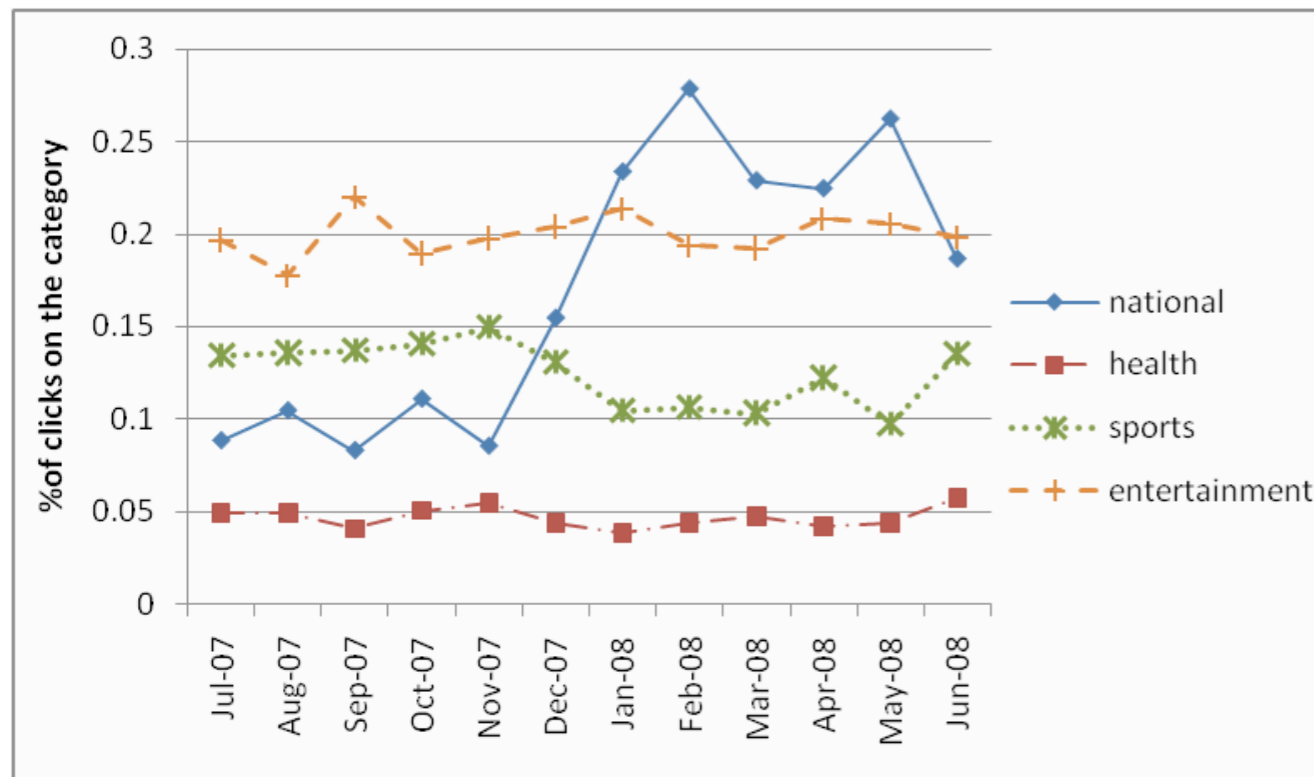


Figure 2. Interest distribution of US users over time

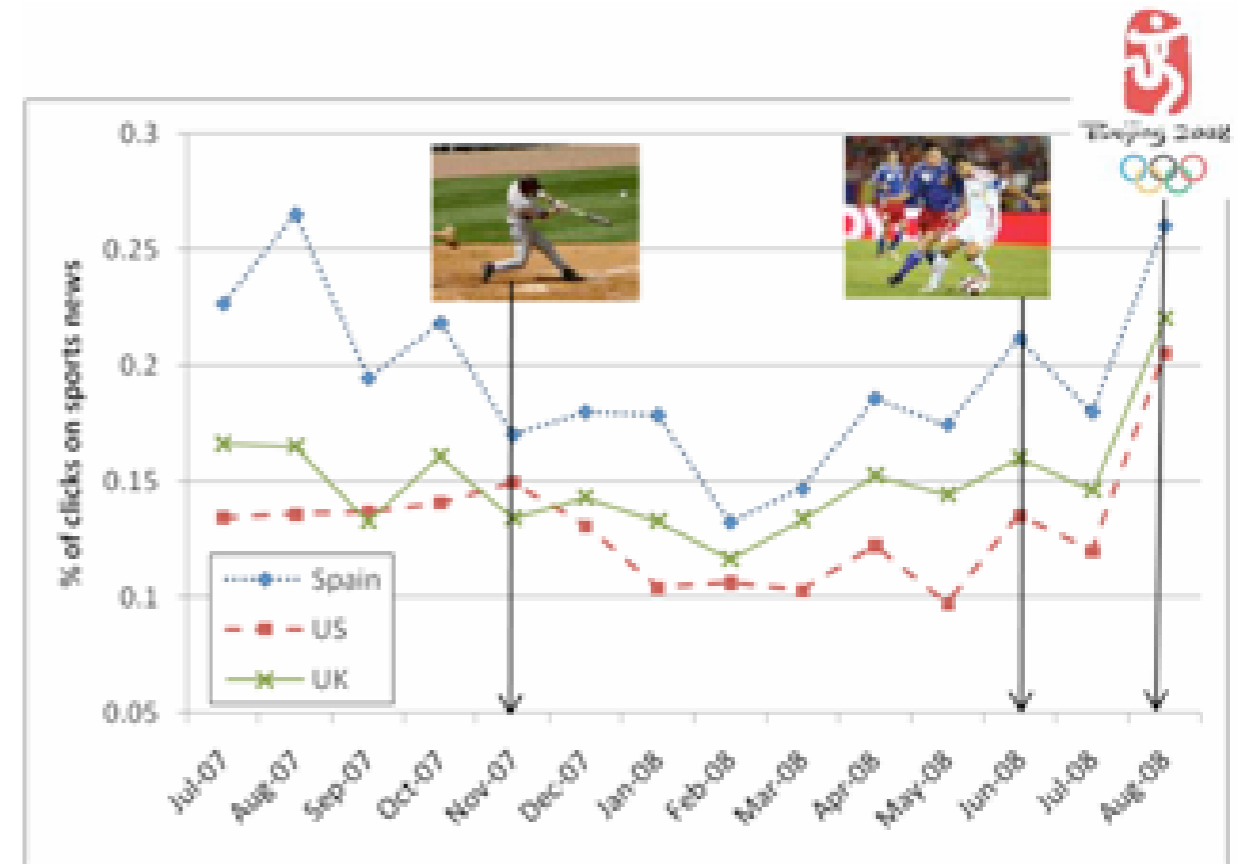


Figure 3. Change of interests in sports news over time



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

- Observations

- The news interests of individual users do **change over time**
- The click distributions of the general public reflect the news trend, which correspond to the **big news events**
- There exists **different news trends in different locations**
- To a certain extent, the individual user's news interests **correspond with the news trend in the location** that the users belongs to



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

- Bayesian Framework for User Interest Prediction

- Predicting user's genuine news interest

- For a specific time period t in the past, the genuine interest of a user in topic category c_i is modeled as

$$p^t(\text{click} \mid \text{category} = c_i)$$

- Using Bayesian rule

$$\begin{aligned} \text{interest}^t(\text{category} = c_i) &= p^t(\text{click} \mid \text{category} = c_i) \\ &= \frac{p^t(\text{category} = c_i \mid \text{click})p^t(\text{click})}{p^t(\text{category} = c_i)} \end{aligned}$$



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

- Bayesian Framework for User Interest Prediction
 - Combining predictions of past time periods

$$\begin{aligned} \text{interest}(\text{category} = c_i) &= \frac{\sum_t \left(N^t \times \text{interest}^t(\text{category} = c_i) \right)}{\sum_t N^t} \\ &= \frac{\sum_t \left(N^t \times \frac{p^t(\text{category} = c_i | \text{click}) p^t(\text{click})}{p^t(\text{category} = c_i)} \right)}{\sum_t N^t} \end{aligned}$$

N^t is the total number of clicks by the user in time period t

- Assume $p^t(\text{click})$ is a constant, then we get

$$\begin{aligned} \text{interest}(\text{category} = c_i) \\ &= \frac{p(\text{click}) \times \sum_t \left(N^t \times \frac{p^t(\text{category} = c_i | \text{click})}{p^t(\text{category} = c_i)} \right)}{\sum_t N^t} \end{aligned}$$



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

- Bayesian Framework for User Interest Prediction

- Predicting user's current news interest

- Use the click distribution of the general public in a short current time period (e.g. in the past hour), represented as $p^0(\text{category} = c_i)$, by using Bayesian rule:

$$\begin{aligned} & p^0(\text{category} = c_i \mid \text{click}) \\ &= \frac{p^0(\text{click} \mid \text{category} = c_i) p^0(\text{category} = c_i)}{p^0(\text{click})} \end{aligned}$$

- Estimate $p^0(\text{click} \mid \text{category} = c_i)$ with genuine interests $\text{interest}(\text{category} = c_i)$

$$\begin{aligned} & p^0(\text{category} = c_i \mid \text{click}) \\ & \propto \frac{\text{interest}(\text{category} = c_i) p^0(\text{category} = c_i)}{p(\text{click})} \\ & \propto \frac{p^0(\text{category} = c_i) \times \sum_t \left(N^t \times \frac{p^t(\text{category} = c_i \mid \text{click})}{p^t(\text{category} = c_i)} \right)}{\sum_t N^t} \end{aligned}$$



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

- Bayesian Framework for User Interest Prediction
 - Predicting user's current news interest
 - Adding a set of virtual clicks G , which is set to be 10 in the system. It can be regarded as a smooth factor.

$$p^0(\text{category} = c_i | \text{click}) \propto \frac{p^0(\text{category} = c_i) \times \left(\sum_t \left(N^t \times \frac{p^t(\text{category} = c_i | \text{click})}{p^t(\text{category} = c_i)} \right) + G \right)}{\sum_t N^t + G}$$



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

- Live traffic experiment
 - Experiments conducted on a fraction (about 10,000 users) of the live traffic at Google News
 - Users were randomly assigned to a control group and a test group. Two groups have the same size
 - Control group uses old recommendation algorithm, while the test group uses the proposed recommendation algorithm



Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

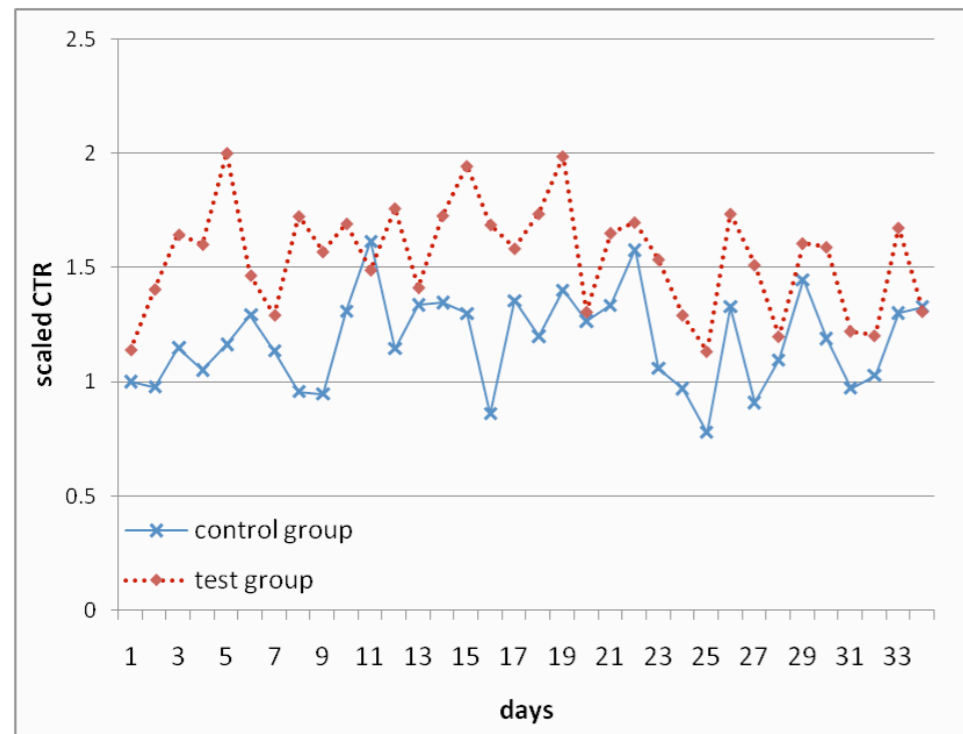


Figure 4. CTR of the recommended news section

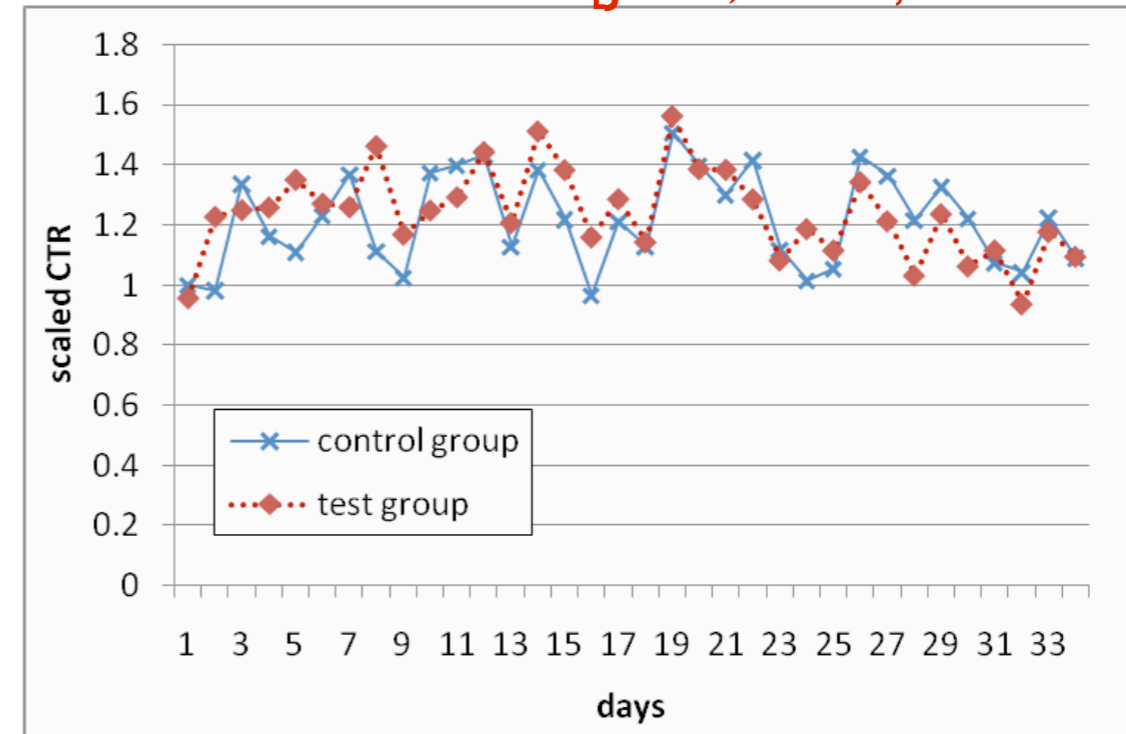


Figure 5. CTR of the Google News homepage

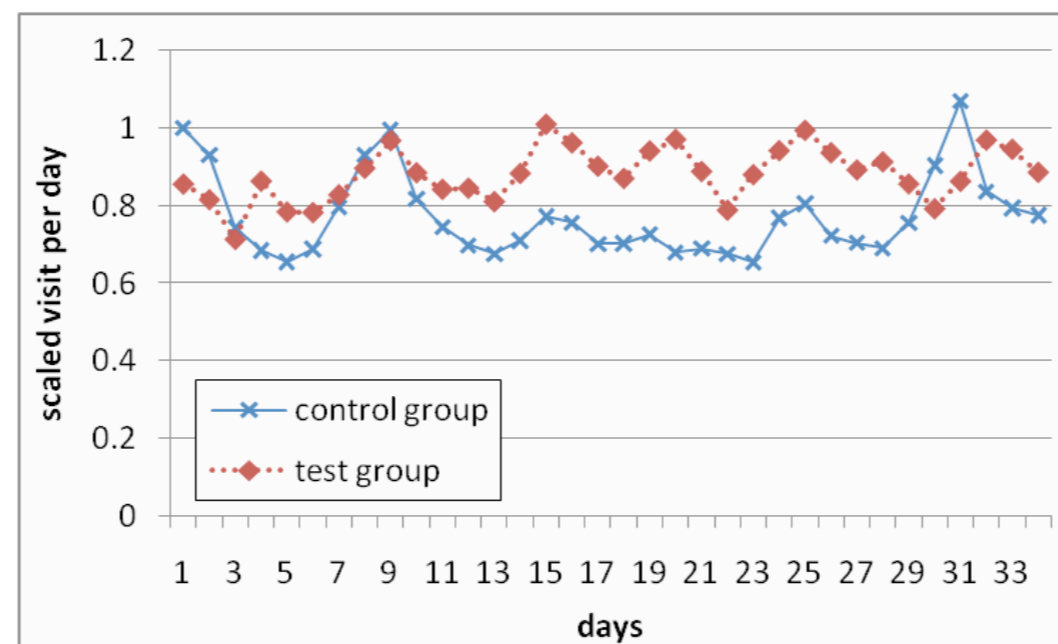


Figure 6. Frequency of website visit per day



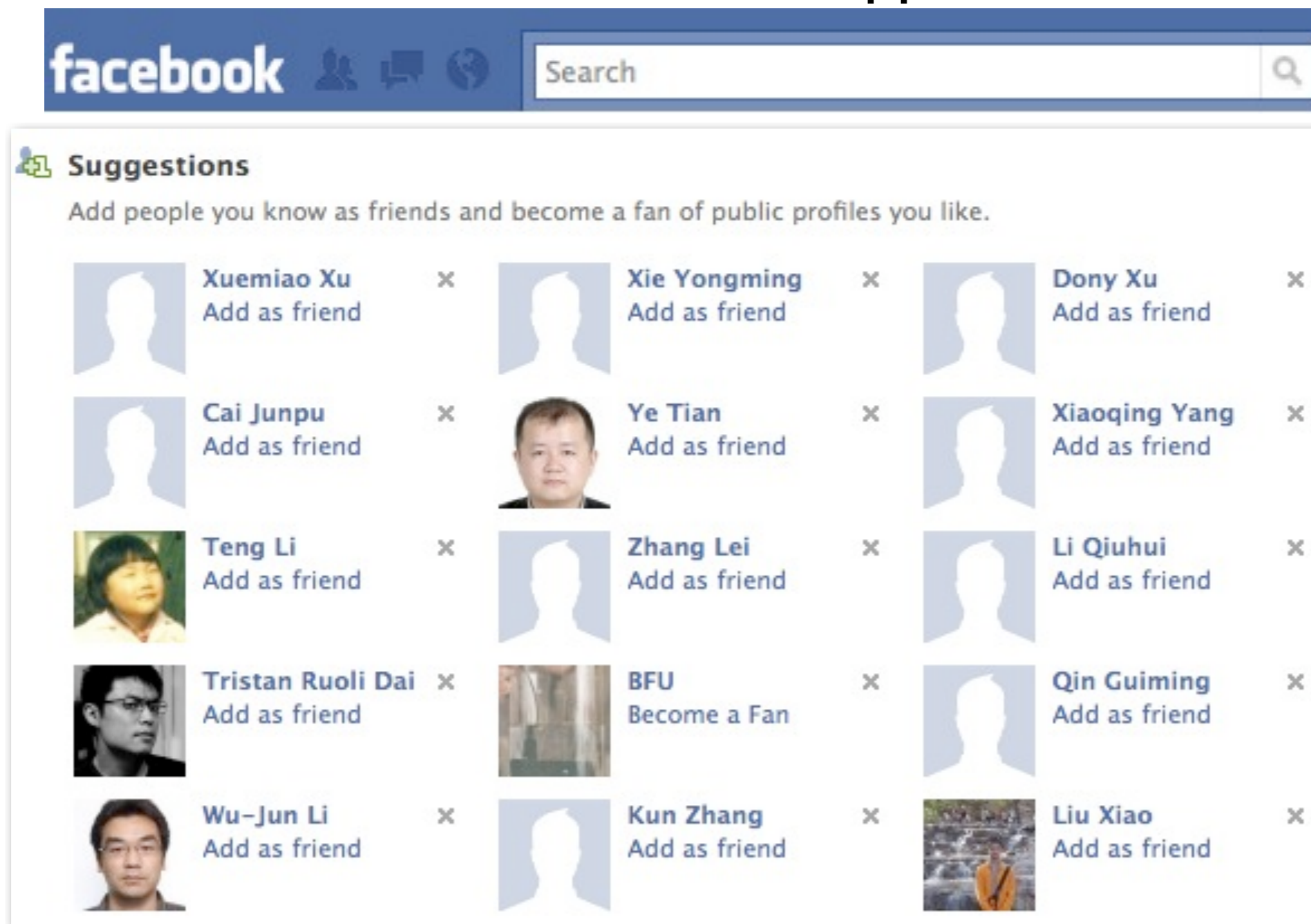
Social Media Analysis

- Social Media Ranking
- Tag Recommendation
- News Recommendation
- **User Recommendation**
- Twitter-powered Recommendation



User Recommendation

- Facebook Service - People You May Know
 - Based on “friend-of-a-friend” approach

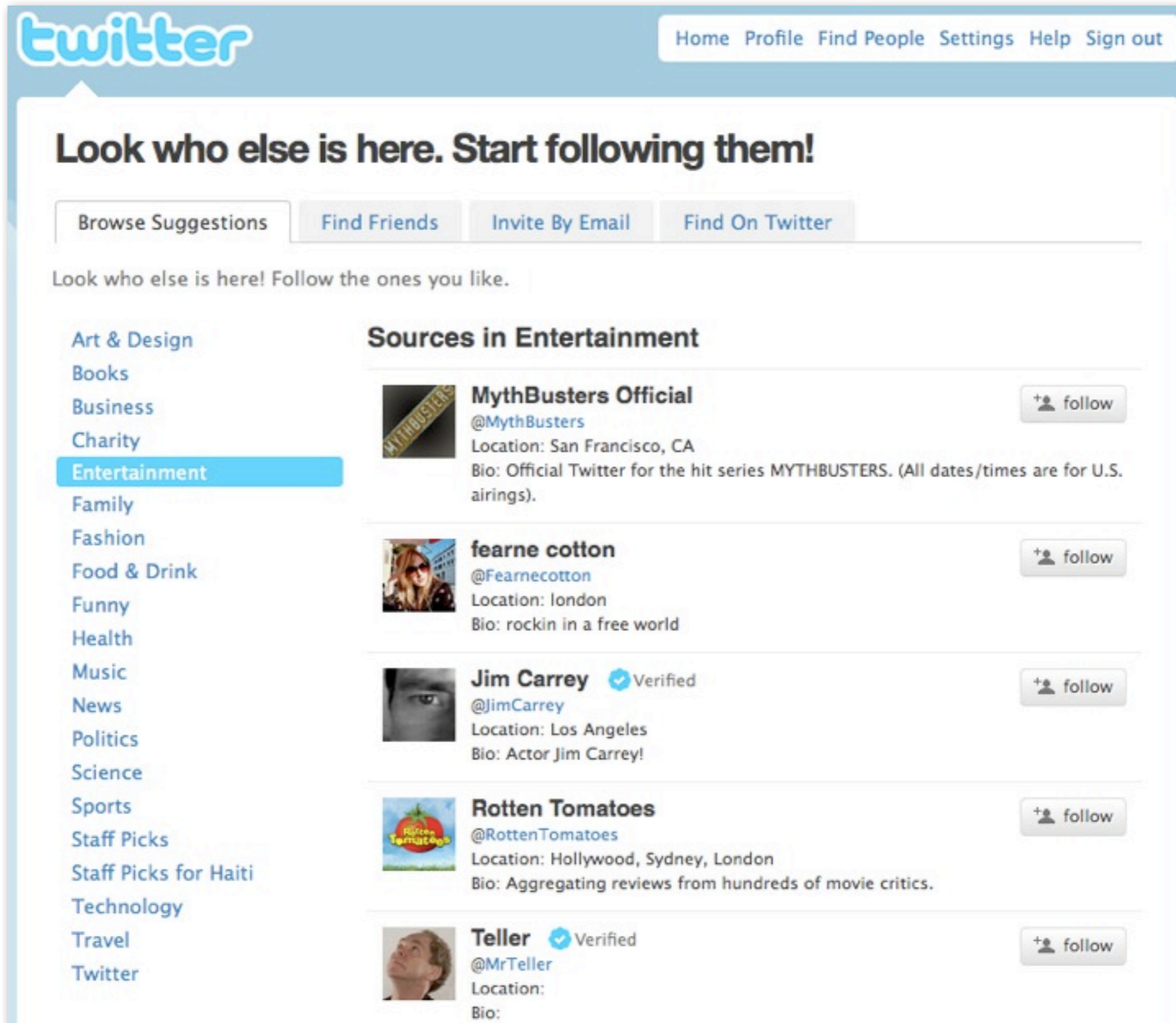


The screenshot shows the Facebook interface with the 'Suggestions' section. The header includes the Facebook logo, navigation icons, and a search bar. Below the header, the 'Suggestions' section is titled 'Suggestions' and includes the text 'Add people you know as friends and become a fan of public profiles you like.' The suggestions are presented in a grid of three columns and five rows. Each suggestion includes a profile picture, the user's name, and a button to either 'Add as friend' or 'Become a Fan'. The users listed are:

Profile Picture	Name	Action
[Placeholder]	Xuemiao Xu	Add as friend
[Placeholder]	Xie Yongming	Add as friend
[Placeholder]	Dony Xu	Add as friend
[Placeholder]	Cai Junpu	Add as friend
[Real Photo]	Ye Tian	Add as friend
[Placeholder]	Xiaoqing Yang	Add as friend
[Real Photo]	Teng Li	Add as friend
[Placeholder]	Zhang Lei	Add as friend
[Placeholder]	Li Qiuhui	Add as friend
[Real Photo]	Tristan Ruoli Dai	Add as friend
[Real Photo]	BFU	Become a Fan
[Placeholder]	Qin Guiming	Add as friend
[Real Photo]	Wu-Jun Li	Add as friend
[Placeholder]	Kun Zhang	Add as friend
[Real Photo]	Liu Xiao	Add as friend



User Recommendation



The screenshot shows the Twitter interface with the 'Look who else is here' section. The top navigation bar includes the Twitter logo and links for Home, Profile, Find People, Settings, Help, and Sign out. Below the navigation bar, the main heading reads 'Look who else is here. Start following them!'. There are four tabs: 'Browse Suggestions', 'Find Friends', 'Invite By Email', and 'Find On Twitter'. A sub-heading says 'Look who else is here! Follow the ones you like.' On the left, a vertical list of categories is shown, with 'Entertainment' highlighted in blue. The main content area is titled 'Sources in Entertainment' and lists five accounts, each with a profile picture, name, handle, location, bio, and a 'follow' button.

Category	Account Name	Handle	Location	Bio	Follow Button
Entertainment	MythBusters Official	@MythBusters	San Francisco, CA	Official Twitter for the hit series MYTHBUSTERS. (All dates/times are for U.S. airings).	+ follow
	fearne cotton	@Fearnecotton	london	rockin in a free world	+ follow
	Jim Carrey	@JimCarrey	Los Angeles	Actor Jim Carrey!	+ follow
	Rotten Tomatoes	@RottenTomatoes	Hollywood, Sydney, London	Aggregating reviews from hundreds of movie critics.	+ follow
	Teller	@MrTeller			+ follow



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

- On social networking sites, people recommendation algorithms are designed to help users:
 - Find known, offline contacts
 - Discover new friends
- Both are challenging problems



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

- Two research questions:
 - How **effective** are different algorithms in recommending people as potential friends?
 - Can a people recommender system effectively **increase** the number of friends a user has?



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

- Test bed
 - Beehive, an enterprise social networking site within IBM
- **Four** different algorithms are tested
- The survey was targeted at a group of 500 users who were asked to answer questions related to their friending behavior



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

- Algorithms

1. Content Matching

- Based on the intuition that “if we both post **content on similar topics**, we might be interested in getting to know each other”
- Based on **TFxIDF** method

2. Content-plus-Link (CplusL)

- Enhances the content matching algorithm with **social link** information derived from social network structure
- Based on the intuition that “By disclosing a network path to a weak tie or unknown person, the recipient will be more likely to accept the recommendation.”



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

- Algorithms

3. Friend-of-Friend (FoF)

- Only leverages social network information of friending
- Based on the intuition that “if many of my friends consider Alice a friend, perhaps Alice could be my friend too”

4. SONAR

- Based on the SONAR system, which aggregates social relationship information from different public data sources within IBM:
(1) Organizational chart; (2) Publication database; (3) Patent database; (4) Friending system; (5) People tagging system; (6) Project wiki; and (7) Blogging system.



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]



Figure 1. Known vs. unknown, Good vs. not good.



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

	Content	CplusL	FoF	SONAR
Content		52.8%	1.8%	8.3%
CplusL			3.3%	9.6%
FoF				13.1%

Table 1. Overlap ratios between recommendations generated by different algorithms.



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

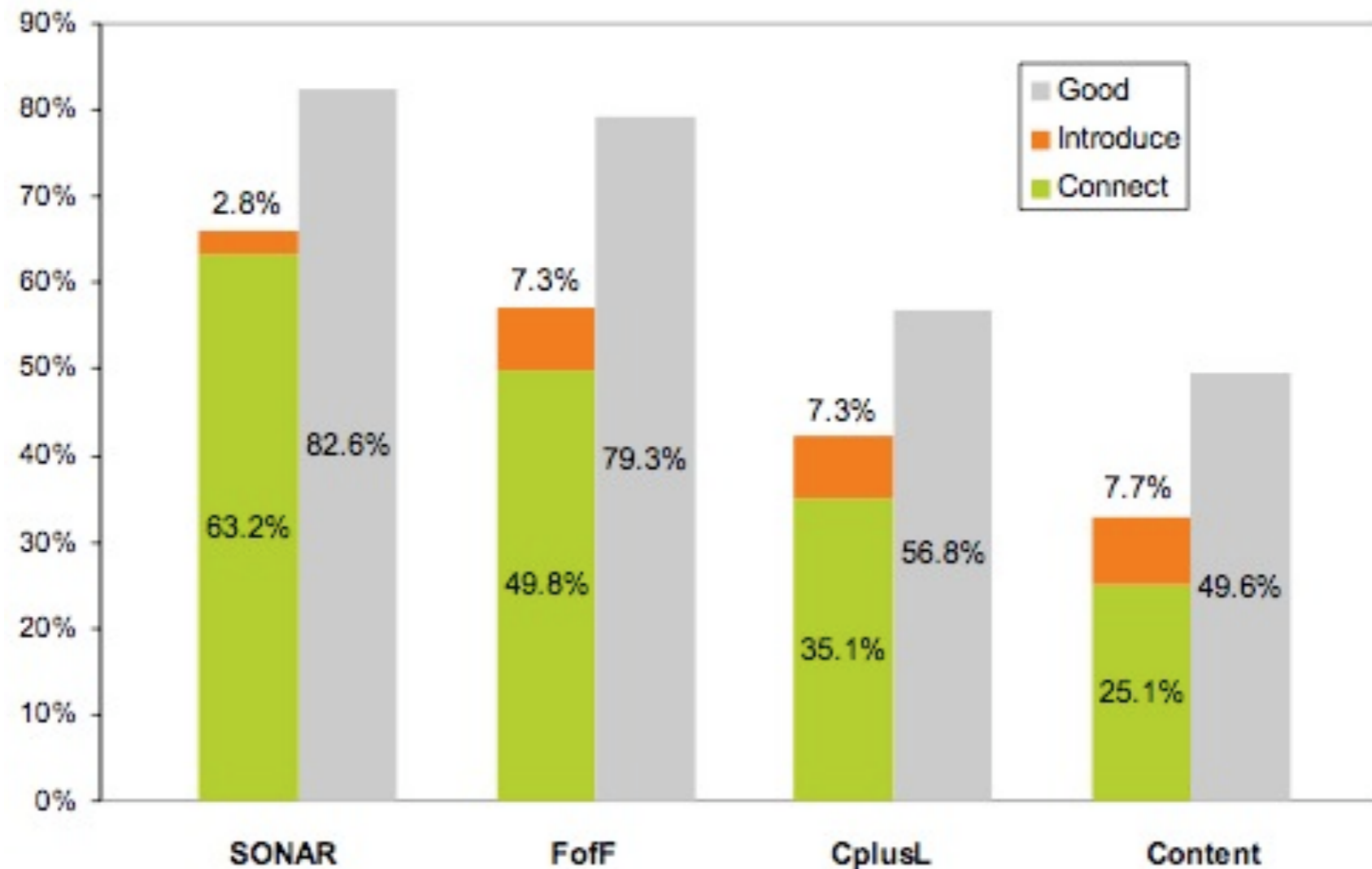


Figure 2. Good recommendations that resulted in actions.



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

SONAR	FoF	CplusL	Content
59.7%	47.7%	40.0%	30.5%

Table 2. Recommendations resulting in connect actions.

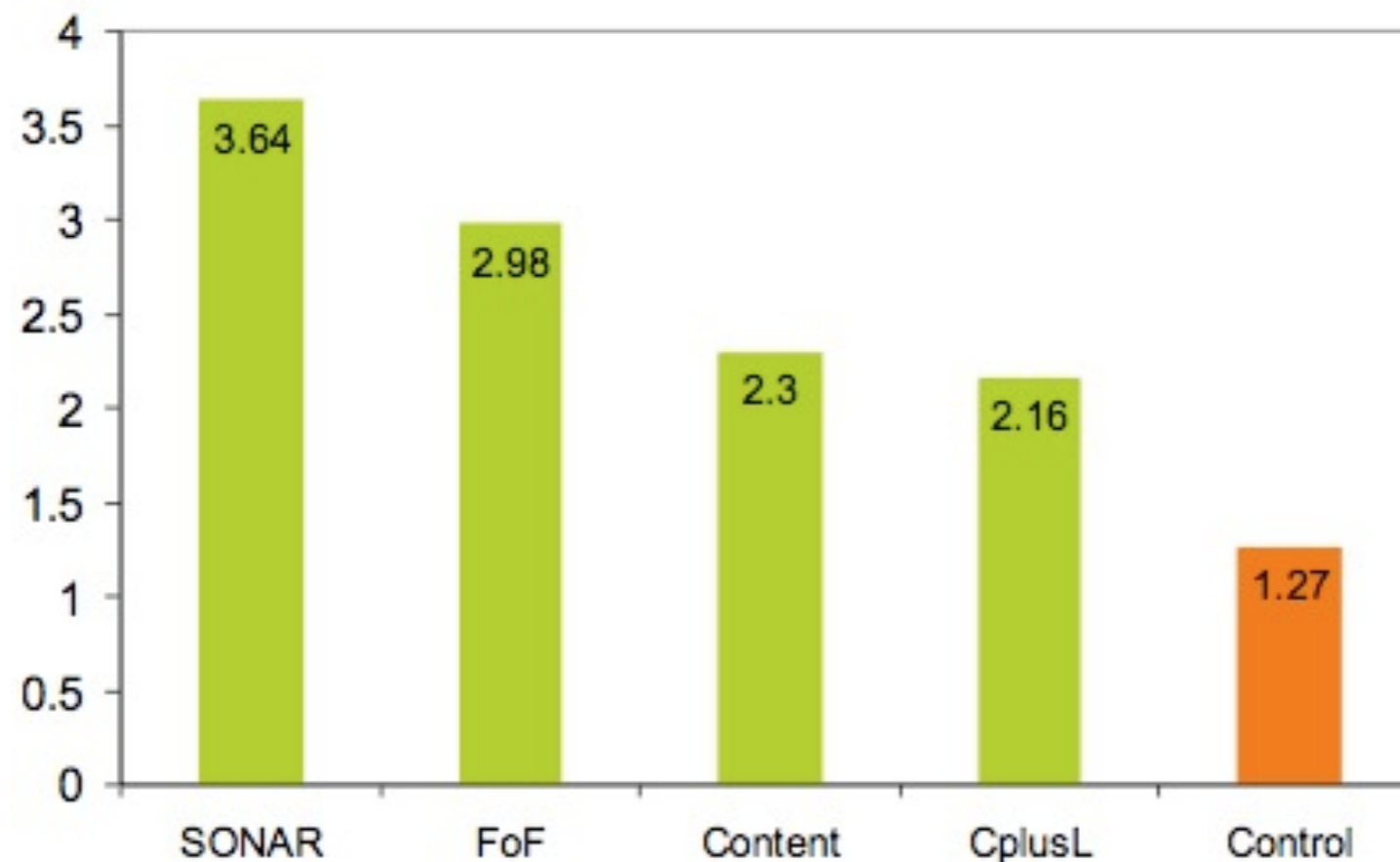


Figure 4. Increase in number of friends.



“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[Jilin Chen, et al., CHI2009]

- Conclusions
 - Relationship based algorithms (FoF and SONAR) outperform content similarity ones (Content and CplusL) in terms of user response
 - Relationship based algorithms are better at finding known contacts whereas content similarity algorithms were stronger at discovering new friends



Social Media Analysis

- Social Media Ranking
- Tag Recommendation
- News Recommendation
- User Recommendation
- **Twitter-powered Recommendation**



Twitter Recommendation Engine

The screenshot shows the Twitter website's recommendation engine. At the top, the Twitter logo is on the left, and navigation links (Home, Profile, Find People, Settings, Help, Sign out) are on the right. Below the logo is a large heading: "Look who else is here. Start following them!". Underneath this heading are four buttons: "Browse Suggestions", "Find Friends", "Invite By Email", and "Find On Twitter". A sub-heading reads: "Look who else is here! Follow the ones you like." On the left side, there is a vertical list of categories: Art & Design, Books, Business, Charity, Entertainment (highlighted in blue), Family, Fashion, Food & Drink, Funny, Health, Music, News, Politics, Science, Sports, Staff Picks, Staff Picks for Haiti, Technology, Travel, and Twitter. The main content area is titled "Sources in Entertainment" and lists five accounts, each with a profile picture, name, handle, location, bio, and a "follow" button:

- MythBusters Official** (@MythBusters) - Location: San Francisco, CA - Bio: Official Twitter for the hit series MYTHBUSTERS. (All dates/times are for U.S. airings).
- fearne cotton** (@Fearnecotton) - Location: london - Bio: rockin in a free world
- Jim Carrey** (Verified) (@JimCarrey) - Location: Los Angeles - Bio: Actor Jim Carrey!
- Rotten Tomatoes** (@RottenTomatoes) - Location: Hollywood, Sydney, London - Bio: Aggregating reviews from hundreds of movie critics.
- Teller** (Verified) (@MrTeller) - Location: [blank] - Bio: [blank]



Twitter-powered Recommendation

The screenshot shows the 'tweetmeme v2' website interface. At the top, it says 'Hottest Links on Twitter'. Below this is a navigation bar with categories: Home, Channels, Comedy, Entertainment, Gaming, Lifestyle, Science, Sports, Technology, World & Bu. A secondary bar below that has 'All', News, Images, and Videos. The main content area is titled 'Everything' and has sorting options: 'Most Recent', 'Top in 24 Hours', and 'Top in 7 Days'. There are three tweets displayed:

- Tweet 1:** 2 tweets. Title: [#KhromeLoungeTonite 857 Washington N Waverly BK,NY after work specials n after party!!Hosted by #DieRich Ent.\(RT\) | TweetPhoto](#). Content: TWEETPHOTO.COM - #KhromeLoungeTonite 857 Washington N Waverly BK,NY after work specials n after (cont) http://tl.gd/romcp. User: PrettyDaGoddd. 0 Comments. Report. Made Popular 27 mins ago.
- Tweet 2:** 150 tweets. Title: [Twitter inicia hoje sua plataforma de publicidade: Promoted Tweets « Brainstorm #9](#). Content: WWW.BRAINSTORM9.COM.BR - Twitter inicia hoje sua plataforma de publicidade: Promoted Tweets. User: cmerigo. 0 Comments. Report. Made Popular 41 mins ago.
- Tweet 3:** 44 tweets. Title: [Boy Genius Reviews the Technology in the 2010 Buick LaCrosse](#). Content: BOYGENIUSREPORT.COM - As part of BGR coverage of NY Auto Show, here's a quick look at the 2010 Buick LaCrosse from a technology perspective. Includes a BGR BUICK logo.



Twitter-powered Recommendation

The screenshot displays the TagWalk website interface, which provides a "sneaky peek into twitter". The main header features the "tag walk" logo with a blue bird icon and the tagline "taking a sneaky peek into twitter". A search bar and social sharing buttons (Twitter, Facebook, LinkedIn, RSS) are located in the top right corner.

TagWalk Stats
Stats about English:
57M tweets, 10.4% retweets, 34.3% with links
577K hashtags, 6.4M talkers, 3.1M to users, 973K web sites
Based on 57M tweets by 6.4M talkers
Last Updated: 2 days ago

Related Hashtags
HashTags related to English:
#jobs #tcot #followfriday #ff #fb #job #iranelection #p2 #hhhs #teaparty #news #quote #lastfm #TweetMyJOBS #hiring #swineflu #php #wordpress #seo #sgp #GOP #tlot #mw2 #fail #Iran #iphone #freelance #photog #photography #tech #love #pr #musicmonday #nowplaying #design #twitter #Squarespace #h1n1 #debill #web +577K

Popular Pictures in English
A small image showing a list of popular pictures, including one of a large indoor arena.

Related Users
Users mentioned in English:
aplusk Mashable stephenfry tommcfly tweetmeme kevinrose TechCrunch guykawasaki Scobleizer dontrythis DavidArchie ZnaTrainer Drudge_Report guardiantech scottbourne addthis JanSimpson taylorswift13 shanselman MrPeterAndre KimSherrell David_Henrie MissKatiePrice Shoq codinghorror bbcworld DonnieWahlberg justinbieber MCHammer jonasbrothers +3.1M
According to 57M tweets by 6.4M users
Last Updated: 2 days ago

Sponsored
Wholesale Sciphone i9
Dual Sim/QuadBand/3.2" Touch Screen
5pcs/lot, \$350/lot. Free Shipping.
Ads by Google

Words
Words used in tweets:
New up now like all get about good how one as it's No More has love time go LOL got they day know twitter when Don't see today there think need too Great going back Really am off had Who he would Here work its want Thanks make via only +16M

Web Sites
Websites in English:
twitpic.com youtube.com twitter.com getafreelancer.com facebook.com

Popular Links in English
What Digital Economy Bill? #debill
1396 tweets since Wed, 7 April by whatdebill Latest: Sun, 11 April
Discover how much power you have as a UK voter in your constituency
335 tweets since Fri, 9 April by Stevelstall Latest: Sun, 11 April
Statute of Anne - Wikipedia, the free encyclopedia
267 tweets since Sat, 10 April by PiratePartyUK Latest: Sun, 11 April
Debillitated
289 tweets since Wed, 7 April by deburca Latest: Sat, 10 April
<http://i.imgur.com/1pXIO.jpg>
232 tweets since Thu, 8 April by lanhogg Latest: Sat, 10 April
Did My MP Show Up or Not?
202 tweets since Wed, 7 April by steve_e Latest: Sat, 10 April
Digital Economy bill: liveblogging the crucial third reading | Technol...
149 tweets since Wed, 7 April by rehagercek Latest: Sun, 11 April
Tumbled Logic - An Open Letter to Siôn Simon, Pete Wishart, David Lamm...
158 tweets since Wed, 7 April by jot Latest: Fri, 9 April
Digital Economy Bill - it's a wash up | The TalkTalk Blog
126 tweets since Thu, 8 April by TalkTalkTips Latest: Sat, 10 April
Daring Fireball: New iPhone Developer Agreement Bans the Use of



Twitter-powered Recommendation

Who Should i Follow?

Find New Twitter Friends

www2010's Recommendations

Tweet This!

Not the results you wanted? Find friends that are:

Less Popular More Popular

Anywhere Closer to



BarackObama

Follow

Location: Washington, DC

Bio: 44th President of the United States

Similar to: [Veronica](#), [katrina_](#), [Pogue](#)

[See more users like BarackObama](#)



Jason

Follow

Location: Los Angeles, CA

Bio: I'm a cereal entrepreneur: Founder of Weblogs, Inc., TechCrunch50, Silicon Alley Reporter, Engadget & Mahalo.com

Similar to: [Veronica](#), [Scobleizer](#), [TechCrunch](#)

[See more users like Jason](#)

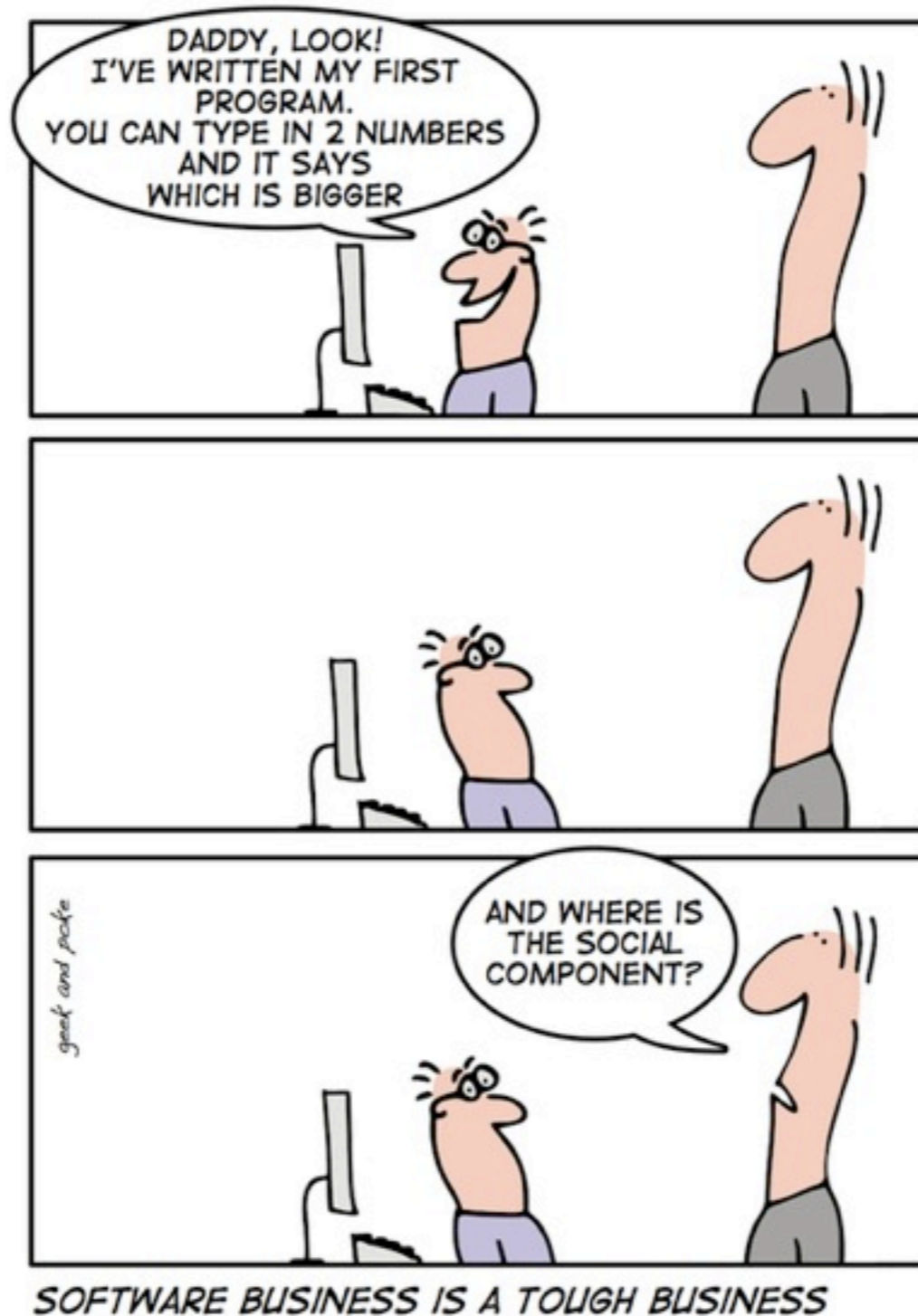


References

- <http://blog.oneriot.com/content/2009/06/oneriot-pulse-rank/>
- <http://uggedal.com/reddit.cf.algorithm.png>
- <http://www.seopedia.org/tips-tricks/social-media/the-digg-algorithm-unofficial-faq/>
- <http://www.technologyreview.com/web/24353/?a=f>
- B. Sigurbjörnsson, and R. van Zwol. Flickr tag recommendation based on collective knowledge. In WWW, pages 327-336, 2008.
- J. Liu, P. Dolan, and E. R. Pedersen. Personalized news recommendation based on click behavior. In IUI, pages 31-40, 2010.
- J. Chen, W. Geyer, C. Dugan, M. J. Muller, and I. Guy. Make new friends, but keep the old: recommending people on social networking sites. In CHI, pages 201-210, 2009



Are You Social Computing Ready?



Q & A



Introduction to Social Computing, Irwin King, 2010 EII PhD School: Cloud Computing, Service Computing & Social Networks, November 23-27, 2010, Brisbane, Australia

