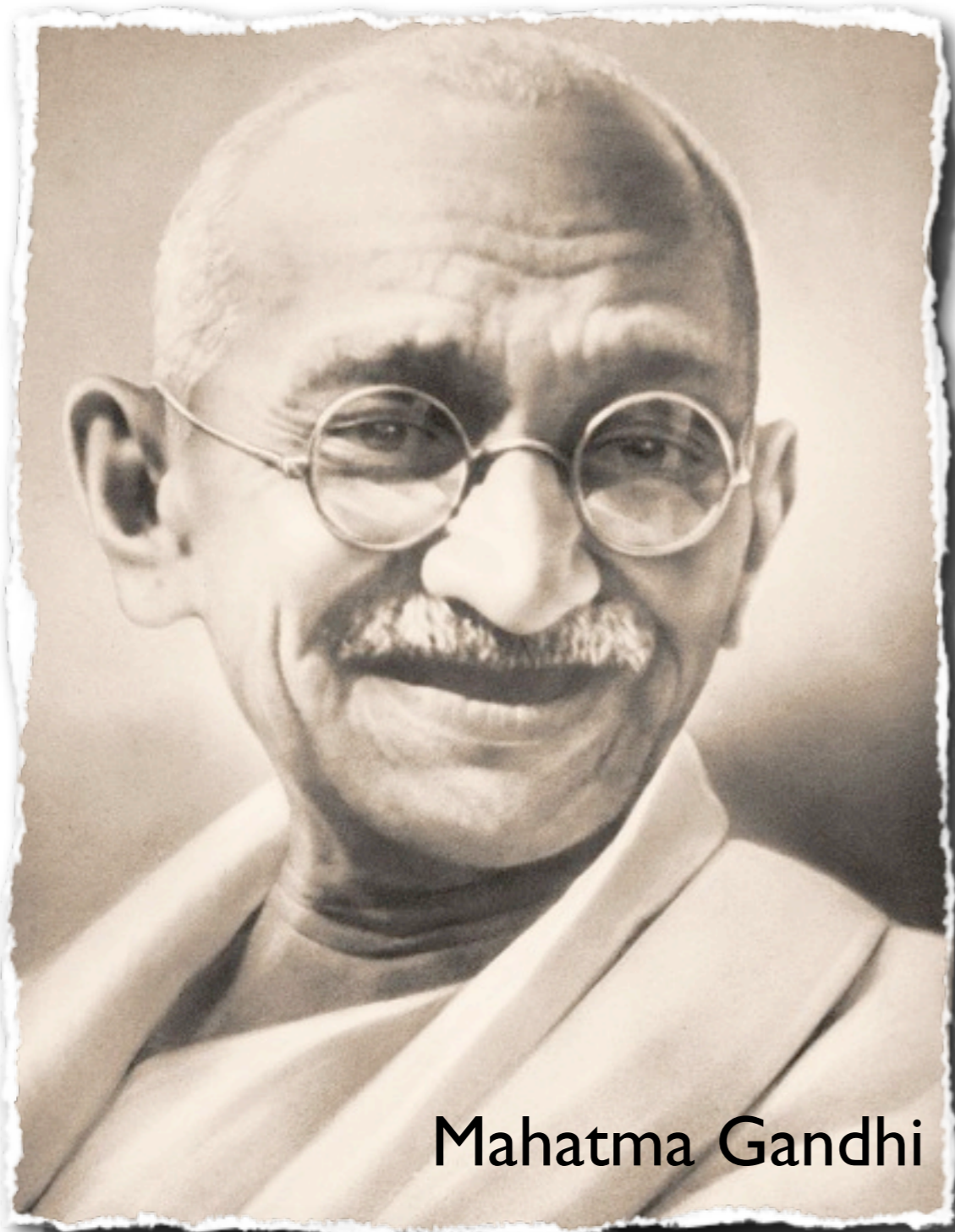


# Introduction to Social Recommendation

Irwin King, Michael R. Lyu, and Hao Ma  
{king, lyu, hma}@cse.cuhk.edu.hk

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





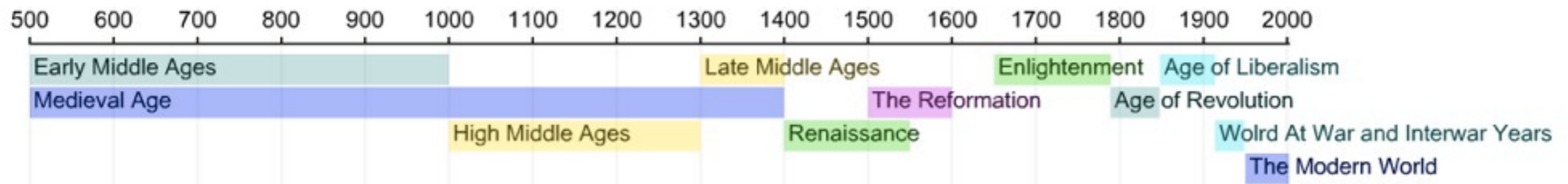
Mahatma Gandhi

*Interdependence is and ought to be as much the ideal of man as self-sufficiency.*

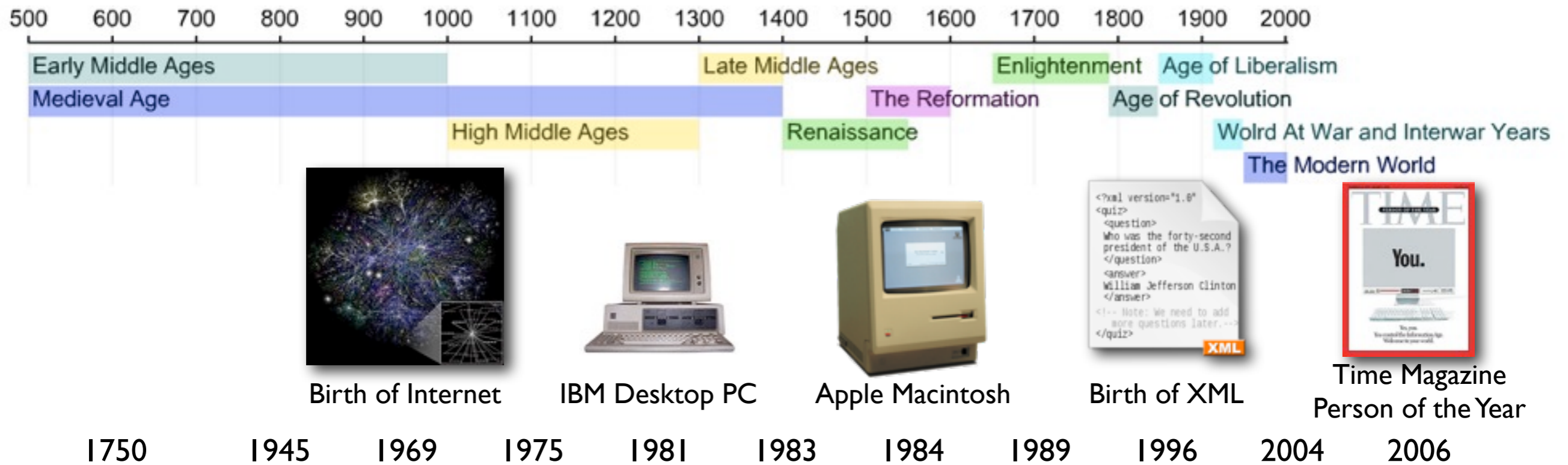
*Man is a social being.*



# A Brief History of the World



# A Brief History of the World



**Industrial  
Revolution**

**Information  
Age**

**Internet  
Age**

**www  
Age**

**Attention  
Age**

ENIAC



The MITS Altair  
Apple II



Time Magazine  
Person of the Year



Birth of WWW



Birth of Web 2.0



Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA







intel.

# revolution in evolution

Highlights from the Journey to 1 Billion PCs



**1971** - Intel, founded by Robert Noyce, Gordon Moore and Andy Grove, introduces the world's first microprocessor and calls it the Intel® 4004.

**1974** - Intel introduces the 8008 microprocessor, which was used in the first commercially successful personal computer - the Intel 8008.

**1976** - Apple Computer, Inc. releases the Apple I, the first single-circuit board computer. The following year, the company introduces the Apple II, the first for a personal computer, the Apple II featured color graphics.

**1977** - Intel introduces the 8080 microprocessor, which was used in the first commercially successful personal computer - the Intel 8080.

**1978** - Intel introduces the 8086 microprocessor, which was used in the first commercially successful personal computer - the Intel 8086.

**1981** - IBM introduces its first personal computer featuring the Intel® 8088 microprocessor. Established the PC revolution and set industry standards that still exist today. The IBM platform enabled hardware makers and software programmers to develop programs and add-on accessories. IBM then, most PCs had been closed and proprietary.

**1982** - Lotus Development Corporation introduces Lotus 1-2-3, which becomes a best-seller application.

**1983** - Apple introduces the Macintosh with a GUI. A GUI is a graphical user interface that provides visual representation for what was previously lines of code. The GUI microprocessor was a 32-bit chip that brought "multi-tasking" capabilities to the PC.

**1984** - Apple introduces the Macintosh with a GUI. A GUI is a graphical user interface that provides visual representation for what was previously lines of code. The GUI microprocessor was a 32-bit chip that brought "multi-tasking" capabilities to the PC.

**1985** - Intel introduces the 386™ microprocessor featuring 275,000 transistors - more than 100 times as many as the original 4004. The 386™ microprocessor was a 32-bit chip that brought "multi-tasking" capabilities to the PC.

**1986** - The number of PCs shipped worldwide reaches nearly 10 million.

**1987** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1988** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1989** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1990** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1991** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1992** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1993** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1994** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1995** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1996** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1997** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1998** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**1999** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2000** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2001** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2002** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2003** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2004** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2005** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2006** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2007** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

**2008** - Intel introduces the Pentium™ processor and Microsoft introduces Windows® 3.1, providing a solid multimedia platform for consumer games and learning applications. Increased processing capabilities, coupled with the availability of affordable CD-ROM drives and sound cards, usher in multimedia on the PC.

For more information, please visit <http://www.intel.com>

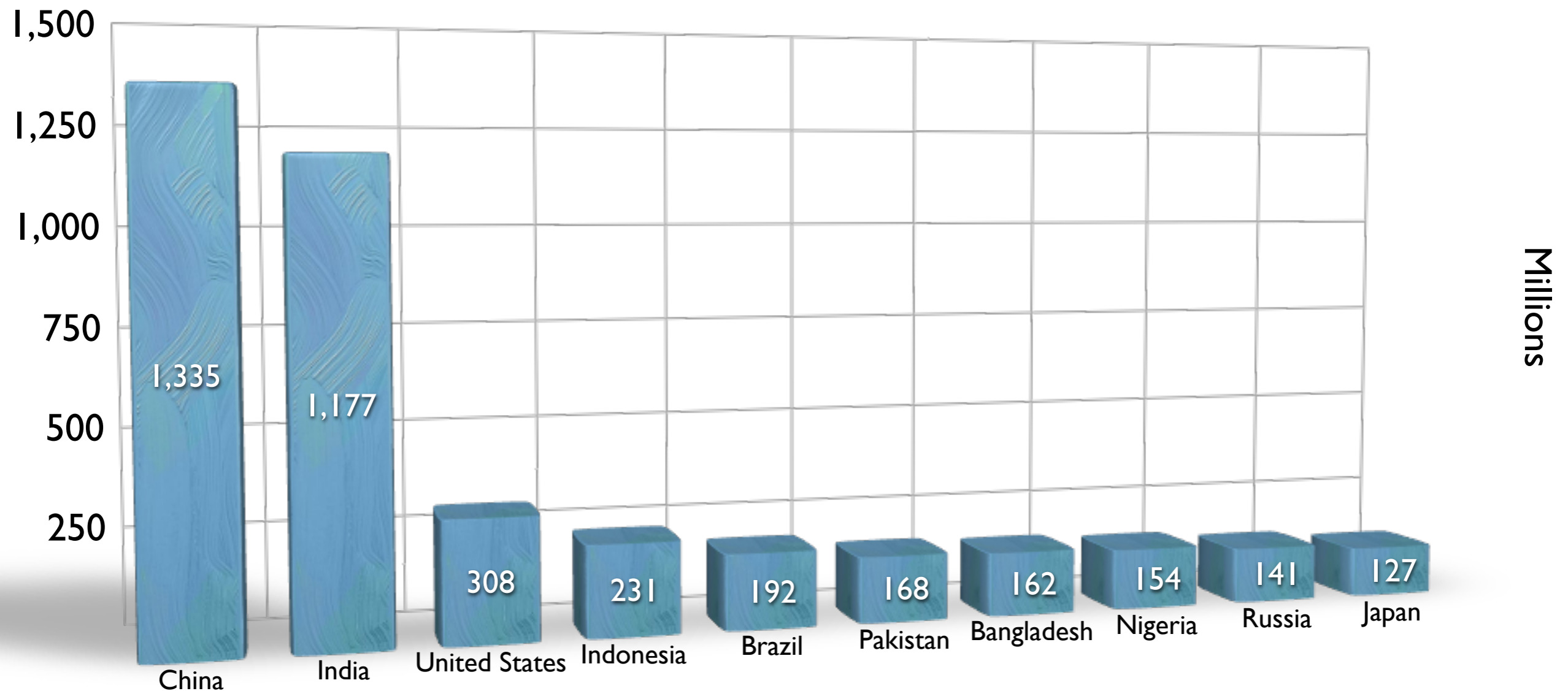
Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA





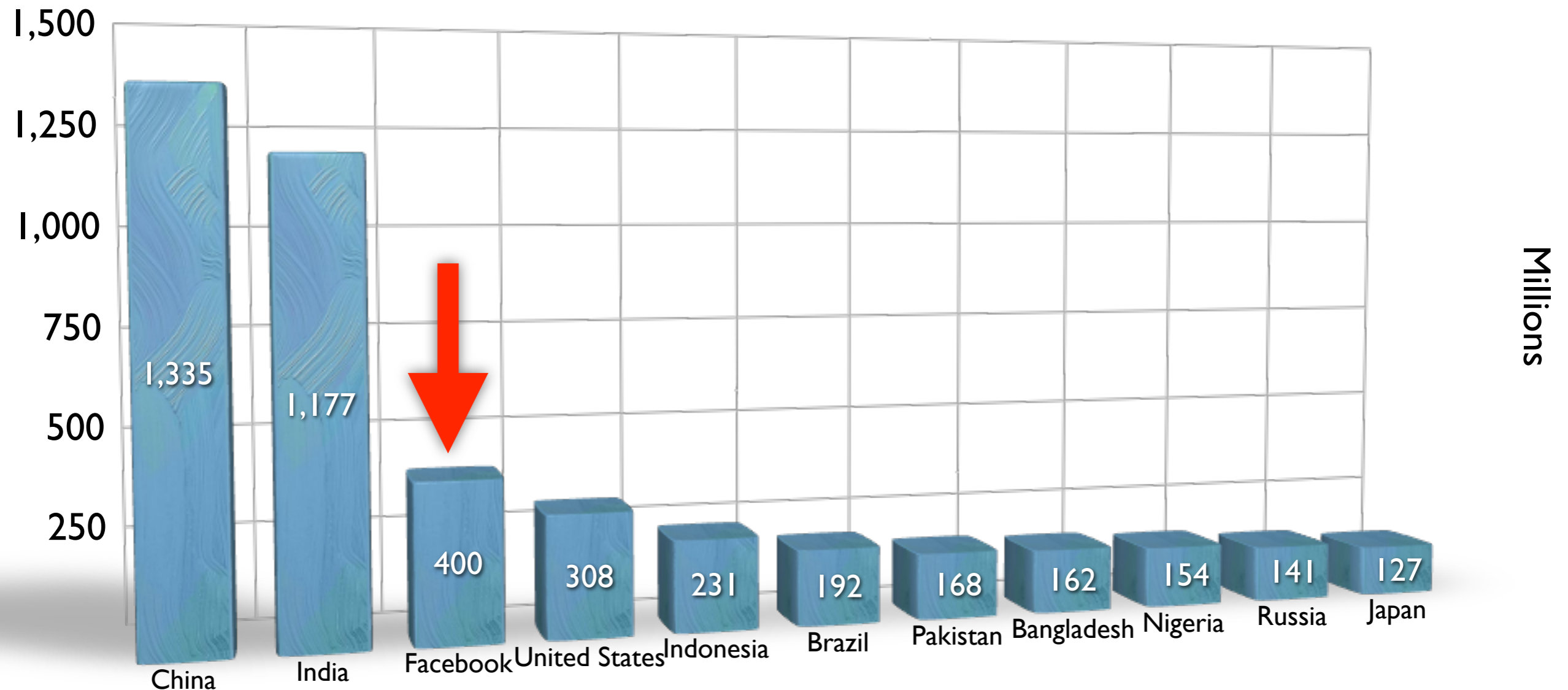
# Top 10 Populations by Countries

as of July 2009



# Top 10 Populations by Countries

as of February 2010



# Facebook's Global Audience

Global Audience: 316,402,840

Data for 11/03/2009



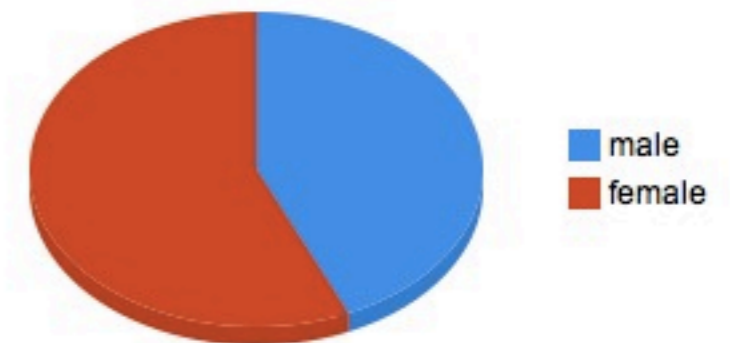
## United States

Country Audience: 94,748,820

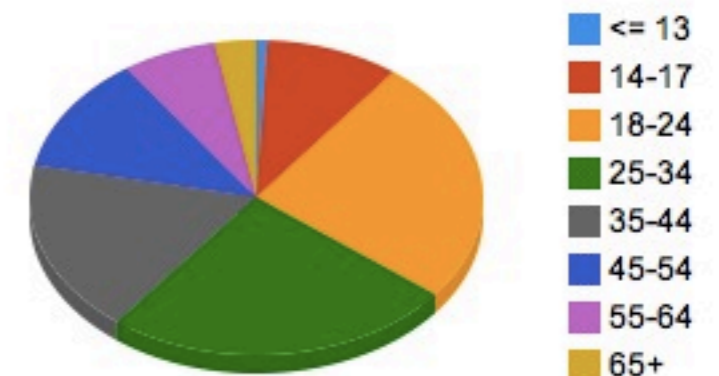
Percent of Global Audience: 29.95%

Share This Site 1543 retweet

United States Male / Female



United States Age Distribution





# Facebook's Growth Stats

## Statistics

### Company Figures

More than 400 million active users  
50% of our active users log on to Facebook in any given day  
More than 35 million users update their status each day  
More than 60 million status updates posted each day  
More than 3 billion photos uploaded to the site each month  
More than 5 billion pieces of content (web links, news stories, blog posts, notes, photo albums, etc.) shared each week

### 10 Largest Countries

1. United States	94,748,820
2. United Kingdom	22,261,080
3. Turkey	14,215,880
4. France	13,396,760
5. Canada	13,228,380
6. Italy	12,581,060
7. Indonesia	11,759,980
8. Spain	7,313,160
9. Australia	7,176,640
10. Philippines	6,991,040

### 10 Fastest Growing Over Past Week

1. Poland	12.46 %	137,900
2. Thailand	10.96 %	161,300
3. Portugal	9.81 %	80,040
4. South Africa	9.25 %	189,080
5. Taiwan	7.82 %	367,400
6. Romania	7.65 %	28,060
7. Germany	7.54 %	350,240
8. Malaysia	7.43 %	236,840
9. Indonesia	6.84 %	752,640
10. Iraq	6.72 %	6,380



# Global Internet Traffic

Alexa as of May 2009	China	USA	Japan	India	Brazil	Global
1	Baidu	Google	Yahoo.jp	Google.in	Google	Google
2	<b>QQ</b>	Yahoo	<b>FC2</b>	Google	<b>Orkut.br</b>	Yahoo
3	Sina	<b>Facebook</b>	Google.jp	Yahoo	Windows Live	<b>YouTube</b>
4	Google.cn	<b>YouTube</b>	<b>YouTube</b>	<b>Orkut.in</b>	Universo Online	<b>Facebook</b>
5	Taobao	<b>Myspace</b>	Rakuten	<b>YouTube</b>	<b>YouTube</b>	Windows Live
6	163	MSN	Livedoor	<b>Blogger</b>	Globo	MSN
7	Google	Windows Live	<b>Ameblo.jp</b>	Rediff	MSN	<b>Wikipedia</b>
8	Sohu	<b>Wikipedia</b>	<b>mixi</b>	<b>Facebook</b>	Google	<b>Blogger</b>
9	Youku	Craigslist	<b>Wikipedia</b>	<b>Wikipedia</b>	Yahoo	Baidu
10	Yahoo	EBay	Google	Windows Live	Terra	<b>Myspace</b>



# The Brave New Words

blogger

wiki

AVATAR

头像

tag cloud

mash-up

unfriend

tweet

blogsphere

twitterati

defriend

hashtags

SEXTING

Folksonomy






# Twitter in Spotlight

HOME PAGE TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

**The New York Times**  
Friday, June 19, 2009

**News**

Search All NYTimes.com   

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS


**The Lede**

[The New York Times News Blog](#)

June 2, 2009, 7:05 PM

## China's Great Firewall Blocks Twitter

By ROBERT MACKEY



Catherine Henriette/Agence France-Presse — Getty Images

Search This Blog

Previous Post: [Bloggers Ponder Last Message From Missing Jet's Computer](#)

Next Post: [Punditry From Bin Laden and Zawahiri on Obama's Trip to the Middle East](#)

**Recent Posts**

June 18 (38 comments) [Latest Updates on Iran's Disputed Election](#)  
To supplement reporting from New York Times correspondents inside Iran on Thursday, The Lede will continue to track the aftermath of Iran's disputed presidential election online.

June 17 (129 comments) [Wednesday: Latest Updates on Iran's Disputed Election](#)  
On Wednesday, The Lede will continue to track the aftermath of Iran's disputed presidential election online, to supplement reporting from New York Times correspondents inside Iran.

June 16 (198 comments) [Tuesday: Latest Updates on Iran's Disputed Election](#)  
To supplement reporting from New York Times correspondents inside Iran, The Lede





# Outline

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



# Introduction

- **Social Platforms**
  - Social Network
  - Social Media
  - Social Games
  - Social bookmarking
  - Social News and Social Knowledge Sharing
- Techniques in Social Recommendation
- Summary



# Web 2.0

- Web as a medium vs. **Web as a platform**
- Read-Only Web vs. **Read-and-Write Web**
- Static vs. **Dynamic**
- Restrictive vs. **Freedom & Empowerment**
- Technology-centric vs. **User-centric**
- Limited vs. **Rich User Experience**
- Individualistic vs. **Group/Collective Behavior**
- Consumer vs. **Producer**
- Transactional vs. **Relational**
- Top-down vs. **Bottom-up**
- People-to-Machine vs. **People-to-People**
- Search & browse vs. **Publish & Subscribe**
- Closed application vs. **Service-oriented Services**
- Functionality vs. **Utility**
- Data vs. **Value**

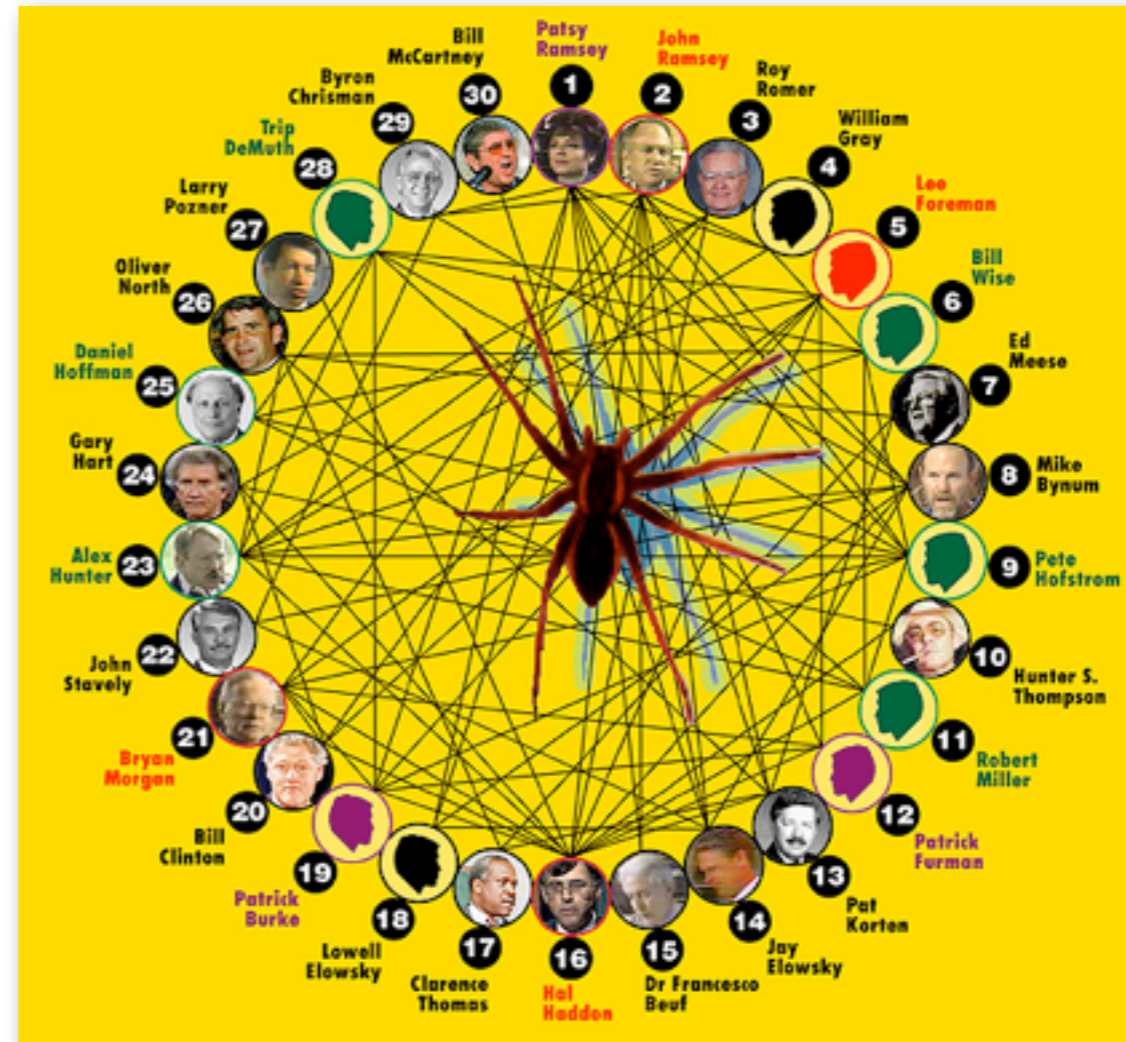


# Social Networks

Society:

**Nodes:** individuals

**Links:** social relationship  
(family/work/friendship/etc.)



S. Milgram and John Guare: **Six Degree of Separation.**  
Social networks: Many **individuals** with diverse **social interactions** between them.





# Milgram's Experiment



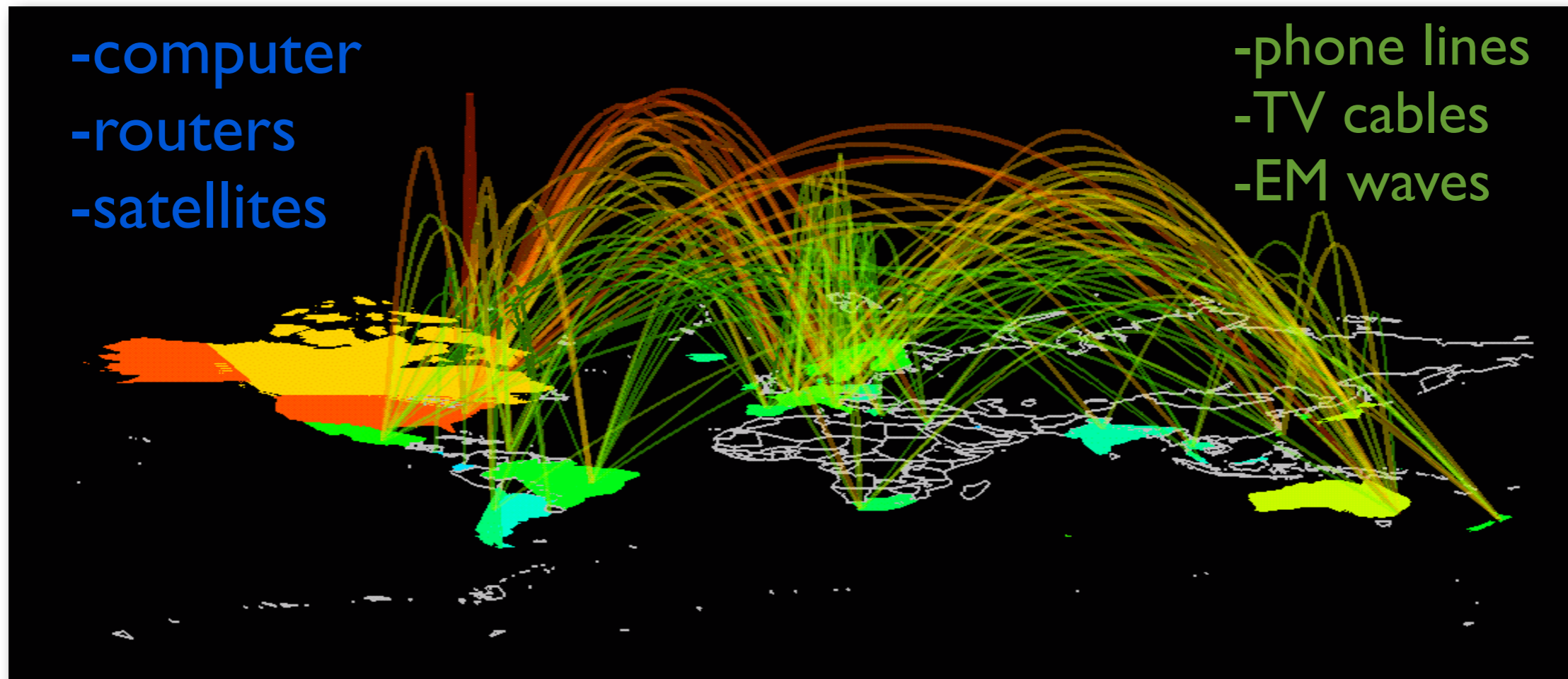
Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA





# Social Networks

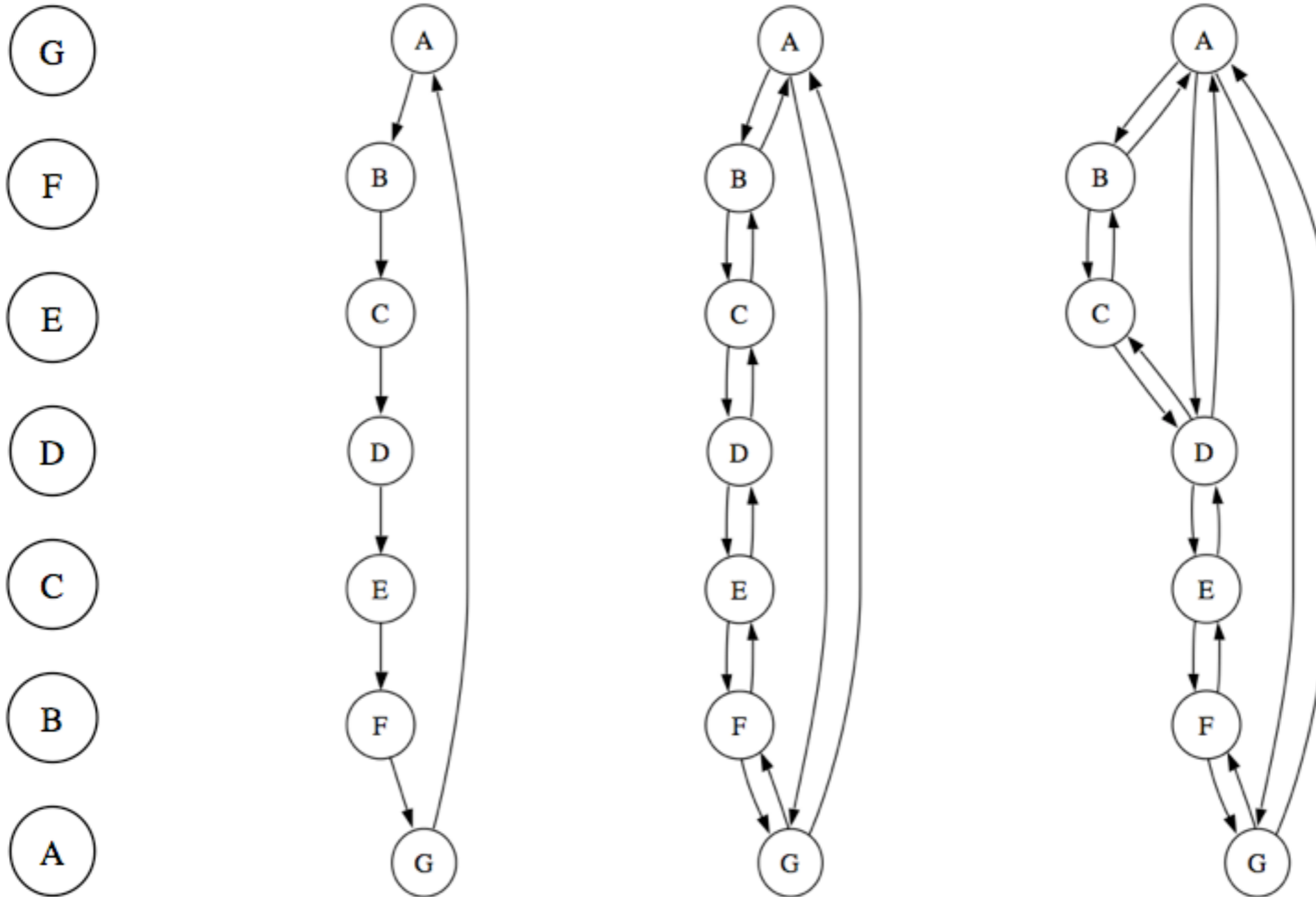
- The Earth is developing an electronic nervous system, a network with diverse **nodes** and **links**.



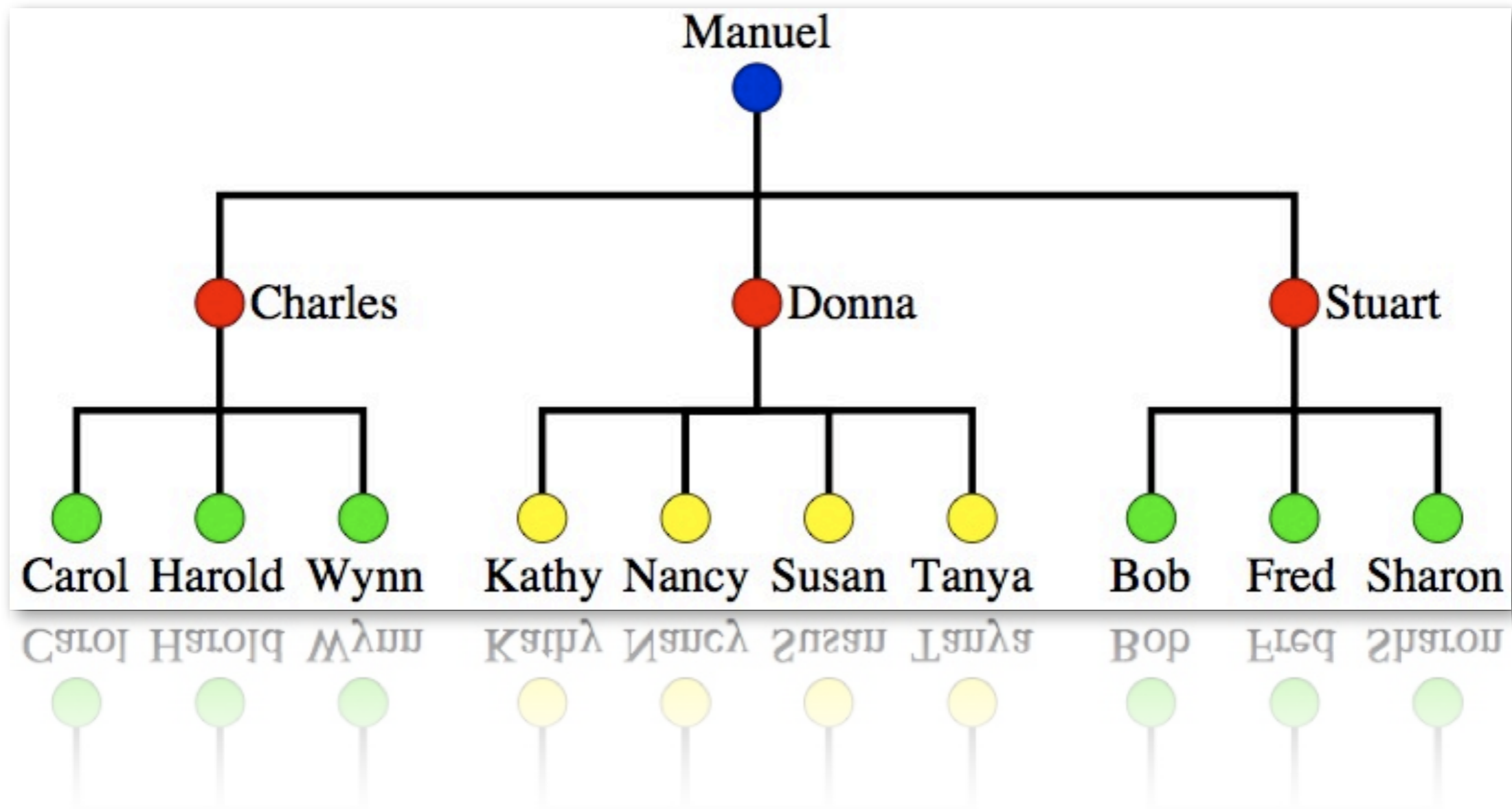
Communication networks: many non-identical components with diverse connections between them.



# The Flow of Information

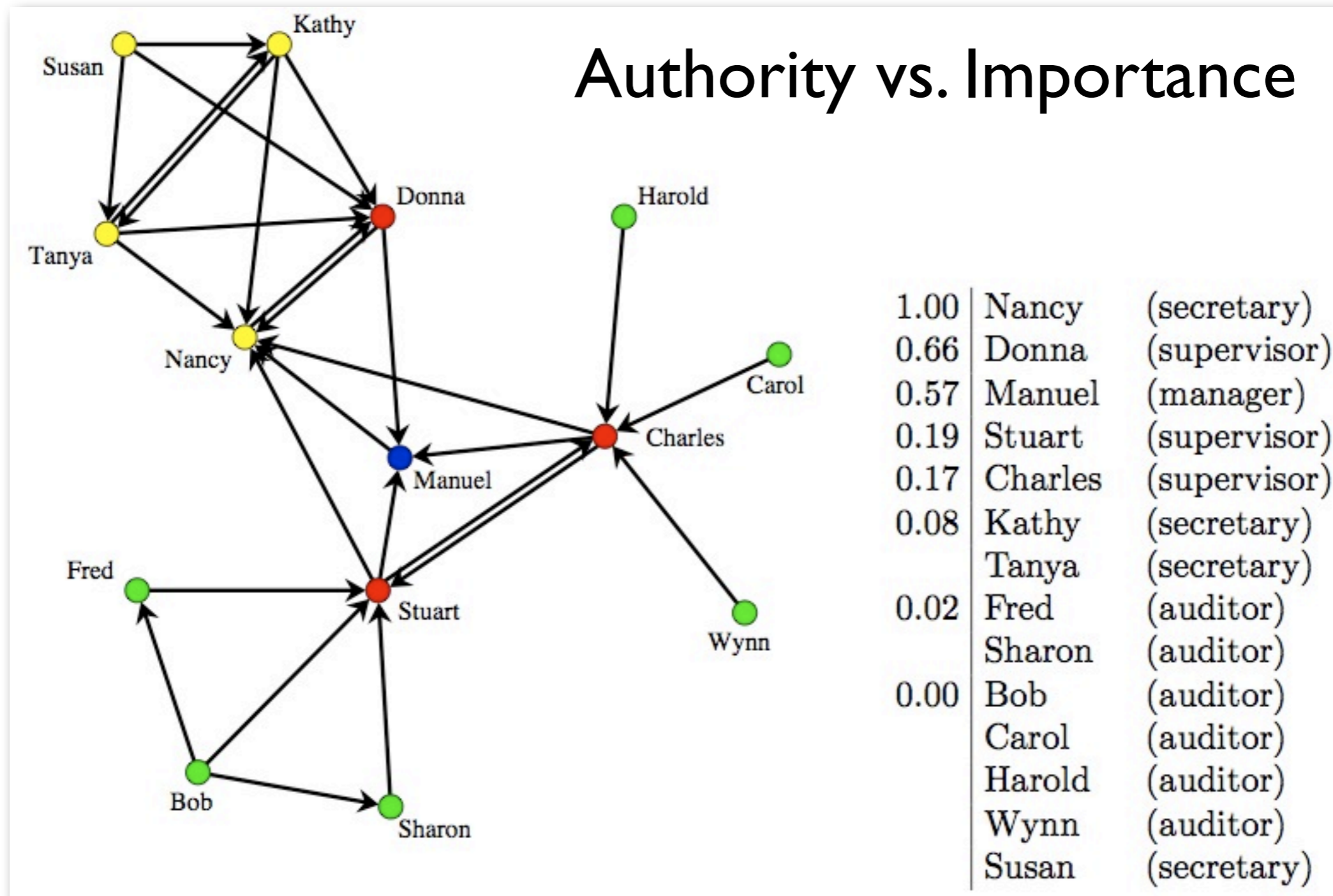


# Organizational Chart





# Social Network Chart





# Social Analytics/Informatics

## Social Informatics

Contact : [Slovenian](#) : [FDV](#)



### Search

[Advanced search](#)

### Login

[New user](#) [Lost password](#)

[SOCIAL INFORMATICS](#)

[STUDY PROGRAMS](#)

[RESEARCH CENTRES](#)

[BIBLIOGRAPHY](#)

### Introduction

- [Concept](#)
- [History](#)

### Relevant Fields

- [Social Informatics](#)
- [Web Content Structure](#)
- [Survey Methodology](#)
- [Marketing Research](#)
- [Social Science Methods](#)
- [Applied Statistics](#)
- [Official Statistics](#)
- [Data Collection](#)
- [Library Science](#)
- [Information Society](#)
- [HC Interaction](#)
- [Information Systems](#)
- [Social ICT Applications](#)
- [Data Modeling & Simulations](#)
- [Media & Communication](#)
- [Science & Technology](#)
- [Arts & Informatics](#)

The notion of social informatics relates to the interaction between society and ICT (information-communication technologies). In its broadest sense it covers:

1. the social consequences of ICT at micro (e.g. social aspects of ICT applications at personal and organisational level) as well as at macro level (e.g. information society studies);
2. the application of ICT in the area of social sciences and social/public sector;
3. the use of ICT as a tool for studying social phenomena (within social science methodology).

Graphical presentation is [here>>](#)

### News

- 07.12.09 [Information Society Free Virtual Library](#)
- 02.12.09 [Job offer: Professor in Social Informatics](#)
- 01.12.09 [Call for papers to "New technologies and data collection in social sciences"](#)
- 09.11.09 [Call for Papers "IASSIST 2010"](#)
- 27.10.09 [Job offer: Associate Professor Position - Department of Social Informatics](#)

[archive](#)

### Blogs

- [Social Informatics by Michael Tyworth](#)
- [Social Informatics - a knol by Per Arne Godejord](#)
- [Pixelcharmer Field Notes: Social Informatics](#)
- [Journal of Social Informatics Blog](#)
- [Social Informatic - International Blog](#)

[more](#)

### Associations

- [The European Survey Research Association](#)
- [Council of American Survey Research Organizations \(CASRO\)](#)
- [Marketing Research Association](#)
- [International Communications](#)

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA






# Politics

HOME PAGE TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

**The New York Times**  
Friday, June 19, 2009

**News**

Search All NYTimes.com   

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS


**The Lede**

[The New York Times News Blog](#)

June 2, 2009, 7:05 PM

## China's Great Firewall Blocks Twitter

By ROBERT MACKEY



Catherine Henriette/Agence France-Presse — Getty Images

Search This Blog

Previous Post: [Bloggers Ponder Last Message From Missing Jet's Computer](#)

Next Post: [Punditry From Bin Laden and Zawahiri on Obama's Trip to the Middle East](#)

**Recent Posts**

June 18 (38 comments) [Latest Updates on Iran's Disputed Election](#)  
To supplement reporting from New York Times correspondents inside Iran on Thursday, The Lede will continue to track the aftermath of Iran's disputed presidential election online.

June 17 (129 comments) [Wednesday: Latest Updates on Iran's Disputed Election](#)  
On Wednesday, The Lede will continue to track the aftermath of Iran's disputed presidential election online, to supplement reporting from New York Times correspondents inside Iran.

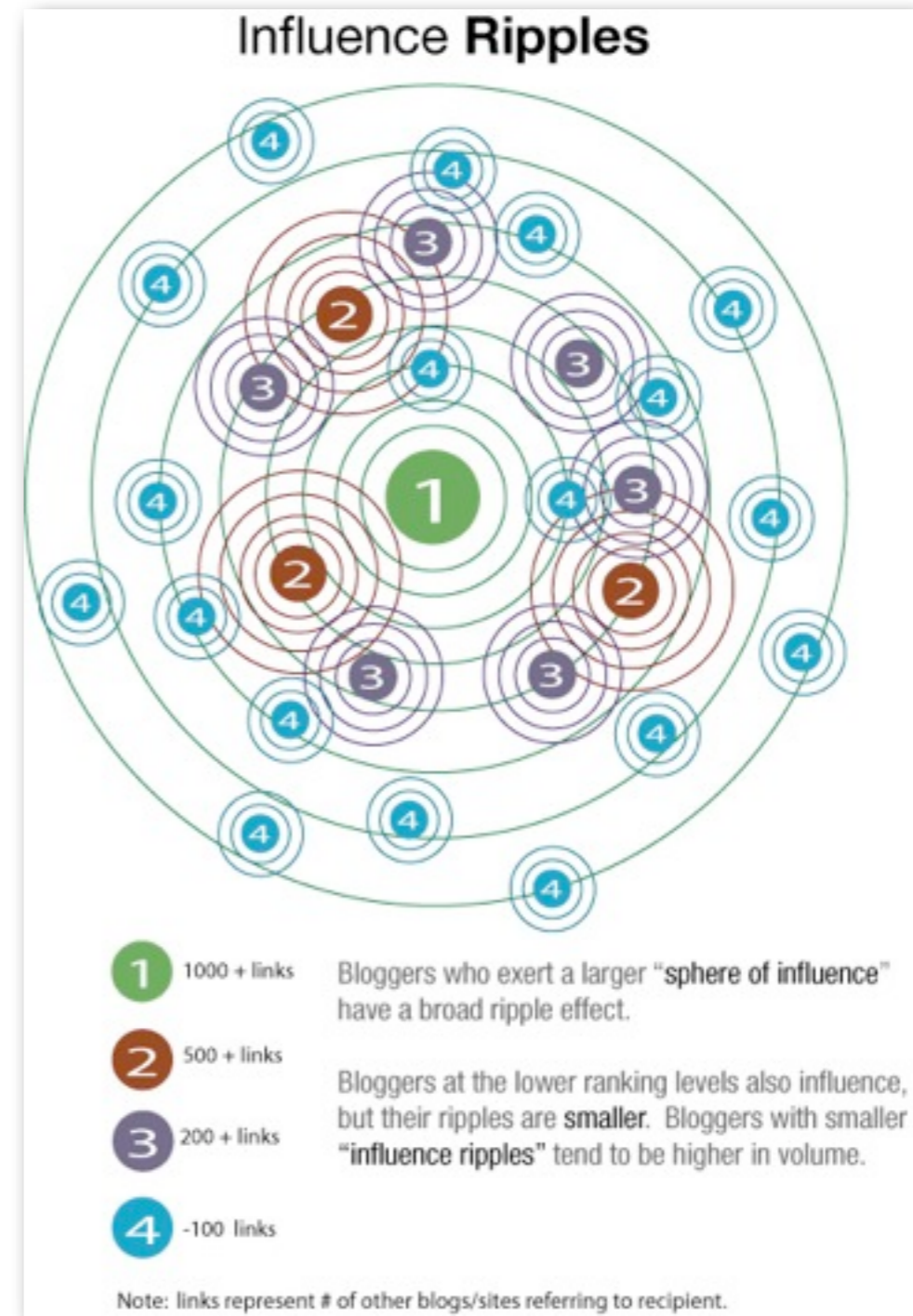
June 16 (198 comments) [Tuesday: Latest Updates on Iran's Disputed Election](#)  
To supplement reporting from New York Times correspondents inside Iran, The Lede





# Commerce

- Social marketing
- Who are the **brokers**?
- Who can exert the **most influence** on buying/selling?
- How **much** should one advertise?



# Public Health

- People's **behavior** can be monitored
- What is on people's mind translates to **search queries**
- Google predicts flu trends...

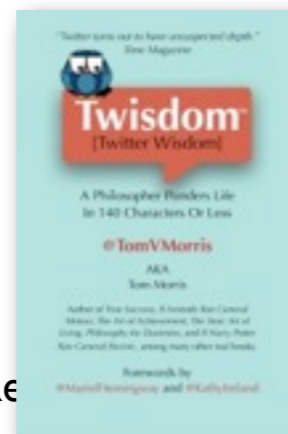
2007–2008 U.S. Flu Activity - Mid-Atlantic Region

ILI percentage



# Twitter Pop Culture

- Twisdom: Twitter Wisdom
  - A Philosopher Ponders Life in 140 Characters or Less
    - “I don’t know the key to success, but the key to failure is trying to please everybody.” Bill Cosby Do what you know in your soul is right!
    - It is a miserable state of mind to have few things to desire, and many things to fear. – Francis Bacon
- The Longest Poem In the World-the awesome twitter poem! 956,644 verses this morning and ~4,000 a day!



Introduction to Social Re Irwin King, Micha



igh, USA





# The YouTube Generation

**THE ACADEMY**  
OF MOTION PICTURE ARTS AND SCIENCES

VISIT OSCARS.ORG  
BECOME A FAN  
SIGN UP FOR NEWS

Oscar®  
Oscars's Channel [Subscribe](#) [Uploads](#) [Favorites](#)

**Steve Martin and Alec Baldwin hosting the Oscars®** 61 ratings ★★★★★  
From: Oscars | March 10, 2010 | 312 views  
Steve Martin and Alec Baldwin, co-hosts of the 82nd Academy Awards®, in their opening monologue.

[View comments, related videos, and more](#)

**Opening Number at the 2010 Oscars®**  
303 views - 4 hours ago

**"The Hurt Locker" winning Best Picture**  
303 views - 4 hours ago

**John Hughes Tribute at the Oscars®**  
301 views - 5 hours ago

**Kathryn Bigelow winning the Oscar® for Directing**  
301 views - 5 hours ago

**Sandra Bullock winning Best Actress**  
309 views - 5 hours ago

**Jeff Bridges winning Best Actor**  
334 views - 5 hours ago

**Steve Martin and Alec Baldwin hosting the**  
312 views - 6 hours ago

**Editing Oscar® Nominees**  
27,246 views - 4 days ago





# The Age of FaceBook

The screenshot shows the Facebook interface for Barack Obama's page. At the top, the Facebook logo and navigation links (Home, Profile, Account) are visible. The page header includes a search bar and a 'Become a Fan' button. The main content area features a profile picture of Barack Obama, followed by tabs for Wall, Info, Boxes, Events, Notes, and Photos. Below these are buttons for 'Barack Obama + Fans', 'Barack Obama', and 'Just Fans'. The main feed contains several posts:

- A post with a large number '8' and the text: "Barack Obama 8: the number of people every minute who are denied coverage, charged a higher rate, or otherwise discriminated against because of a pre-existing condition." It includes a link to "Health Reform by the Numbers: 8" on www.whitehouse.gov and is dated "27 minutes ago".
- A post titled "Barack Obama Speaking about health insurance reform this morning at Arcadia University - starting at 11:00 a.m. ET." It features a "LIVE" video player and a link to "President Obama Speaks on Health Insurance Reform" on www.whitehouse.gov, dated "Yesterday at 12:21am".
- A post with a video player and the text: "Barack Obama I need your help in urging all Americans who want health reform to make their voices heard." It includes a link to "President Obama's message to supporters: 'We need you in this final march for reform'" on www.youtube.com and is dated "March 5 at 8:14am".

On the right side, there is a sidebar with a "Create an Ad" section and a "Connect With More Friends" section featuring an envelope icon and the text: "Share the Facebook experience with more of your friends. Use our simple invite tools to start connecting." Below this is a "More Ads" link.

On the left side, there is a section titled "Add to My Page's Favorites" and "Suggest to Friends". Below that, a text box states: "This page is run by Organizing for America, the grassroots organization for President Obama's agenda for change. To visit the White House Facebook page, go to: http://bit.ly/2bVCm. OFA is a special project of the Democratic National Committee."

At the bottom left, there is an "Information" section with the following details:

- Current Office**
- Office: President of the United States





# Social Networking Sites

- Example of Social Networking Sites: FaceBook, MySpace, Blogger, QQ, etc.





# Social Search

- Social Search Engine
- Leveraging your social networks for searching

The screenshot shows the Eureka! Swicki website. At the top, there's a navigation bar with links for 'build new swicki', 'swicki directory', 'about swickis', and 'about eureka!'. Below this is a large banner with the text 'a custom search portal around the topic of your choice powered by your community'. A search bar is visible on the right. The main content area is divided into several sections: 'Build a swicki!' with a description and a 'Build a swicki' button; 'Eureka! news' with a 'Now out of beta!' announcement and several news items; 'Browse the directory' with a search prompt; and a grid of featured swickis under categories like 'Recently created', 'Top swickis', 'DIY: home improvement swicki showcase', 'Computers', 'Business', 'Home', and 'Regional'.

The screenshot shows the Delver social search engine interface. It features a logo on the left and the text 'delver:: liad agmon edit'. Below the logo, there's a message: 'Your friends are the best source of information! Look for information, media and people within your network'. A search bar with a '(Go)' button is present. The main area displays a network graph with user avatars and connections. A central profile card for 'Nira Rabiner' is highlighted, showing a photo and some details. The interface is clean and modern, with a focus on social connections.





# Social Media

The screenshot shows the YouTube homepage with the following elements:

- Header:** YouTube logo, "Broadcast Yourself™", navigation tabs (Home, Videos, Channels, Community), search bar, and "Upload" button.
- Videos being watched right now...:** A row of five video thumbnails with durations (02:13, 03:29, 01:58, 07:01, 03:53).
- Promoted Videos:** Four video thumbnails with titles like "Think Again Awards" and "第14屆十大電視廣告頒獎典禮 - 飛出...".
- Featured Videos:** A list of featured videos with titles and view counts:
  - David Sedaris delivers a pizza:** From [weaknights](#), Views: 11,313, 5 stars, 01:01. More in [Comedy](#).
  - Erbert and Gerbert's Candle Cannon:** From [candlecannon](#), Views: 109,029, 5 stars, 02:34. More in [Entertainment](#).
  - Girl's Night Out:** From [danidovine](#), Views: 169,435, 5 stars, 03:49. More in [Comedy](#).
  - Lionel Neykov - Freeze My Senses:** From [LionelNeykov](#), Views: 150,758, 5 stars, 03:35. More in [Music](#).
- What's New:** A yellow box containing:
  - YouTube Mobile:** New! Watch ALL YouTube videos on your mobile device.
  - Warp!** Visually fly through YouTube videos in the Fullscreen player.
  - RSS Feeds:** Click on the "RSS this page" link to get fresh videos delivered.
  - SXSW on YouTube:** For the next week and a half, the SXSW festival is taking over Austin, Texas, to celebrate music, film and all things interactive. [Read more in our Blog](#).

The screenshot shows the Flickr homepage with the following elements:

- Header:** Flickr logo, "Sign In", "Create Your Account" button, and a search bar.
- Main Content:** A large photo of a small plant growing in a crack in the pavement, with the text "Share your photos. Watch the world." and a "SEARCH" button.
- Footer:** Four icons representing different features: "Share & stay in touch", "Upload & organize", "Make stuff!", and "Explore...".

The screenshot shows the Second Life homepage with the following elements:

- Header:** Second Life logo, "Your World. Your Imagination.", "Resident, Login | Join", and navigation tabs (What is Second Life?, Showcase, Community, Blog, Support).
- Main Content:** A large image of a man and a woman flying through the sky, with the text "Get Started! Membership is FREE!" and "Second Life is an online, 3D virtual world imagined and created entirely by its Residents."
- Footer:** A section titled "Your Organization in Second Life!" with a "Visit Second Life Now!" button.

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA





# Social News/Mash Up

The Digg website interface features a top navigation bar with 'Join Digg', 'About', and 'Login'. Below this is a search bar and a 'Popular' filter. The main content area is titled 'News, Videos, Images' and displays a list of articles. The first article is 'Microsoft Demos "ADD TO DIGG" Feature in IE8' with 104 votes. Other articles include 'It was only a matter of time, The SIMS 3 Official' (161 votes) and 'Universe submerged in a sea of chilled neutrinos' (151 votes). A sidebar on the right lists 'Top in All Topics' with items like 'The ravages of aging: Sean Connery, 20 years ago vs Today' (2387 votes) and 'A 10-Year Prison Sentence for Selling Light Bulbs' (1718 votes).

The Twitter website's 'What is Twitter?' page includes a navigation menu with 'What?', 'Why?', and 'How?'. A 'Watch a video!' button is present. A sign-in form asks for a 'user name or email address' and a 'password', with a 'Remember me' checkbox and a 'Sign in' button. A green button offers 'Already using Twitter from your phone? Click here.' Below the sign-in form, a text block states: 'Twitter is a service for friends, family, and co-workers to communicate and stay connected through the exchange of quick, frequent answers to one simple question.' A map overlay shows '8 new tweets' and a tweet from 'Killane' saying 'I feel odd' posted '17 minutes ago in North of Seattle'.

The FoxyTunes website features a search bar for 'artist or song name' and a 'Go' button. It displays a profile for 'Björk' with a search bar for 'Albums' and 'Tracks'. The main content area is divided into several sections: 'Videos on YouTube' with 'All is full of love' (4:09) and 'bjork-hunter' (3:38); 'Lyrics from Yahoo! Music' with a list of tracks including '5 Years', 'Alarm Call', and 'Bachelorette'; 'Flickr Photos' with 'Selected Photos' and 'More on Flickr'; and 'Artist on Last.fm' with 'The Sugarcubes' and 'Goldfrapp'. A 'Music on Hype Machine' section is also visible at the bottom.







# Social/Human Computation

Security Check: Enter both words below, separated by a space. What's This?  
Can't read this? Try another.  
[Try an audio captcha](#)



Text in the box:

I have read and agree to the [Terms of Use and Privacy Policy](#)

[Sign Up](#)

[Problems signing up? Check out our help pages](#)

Προβλήματα σημειώνοντας τα κωδικά ασφαλείας? Ελέγξτε τις σελίδες βοήθειας

Security Check: Enter both words below, separated by a space. What's This?  
Can't read this? Try another.  
[Try an audio captcha](#)



Text in the box:

I have read and agree to the [Terms of Use and Privacy Policy](#)

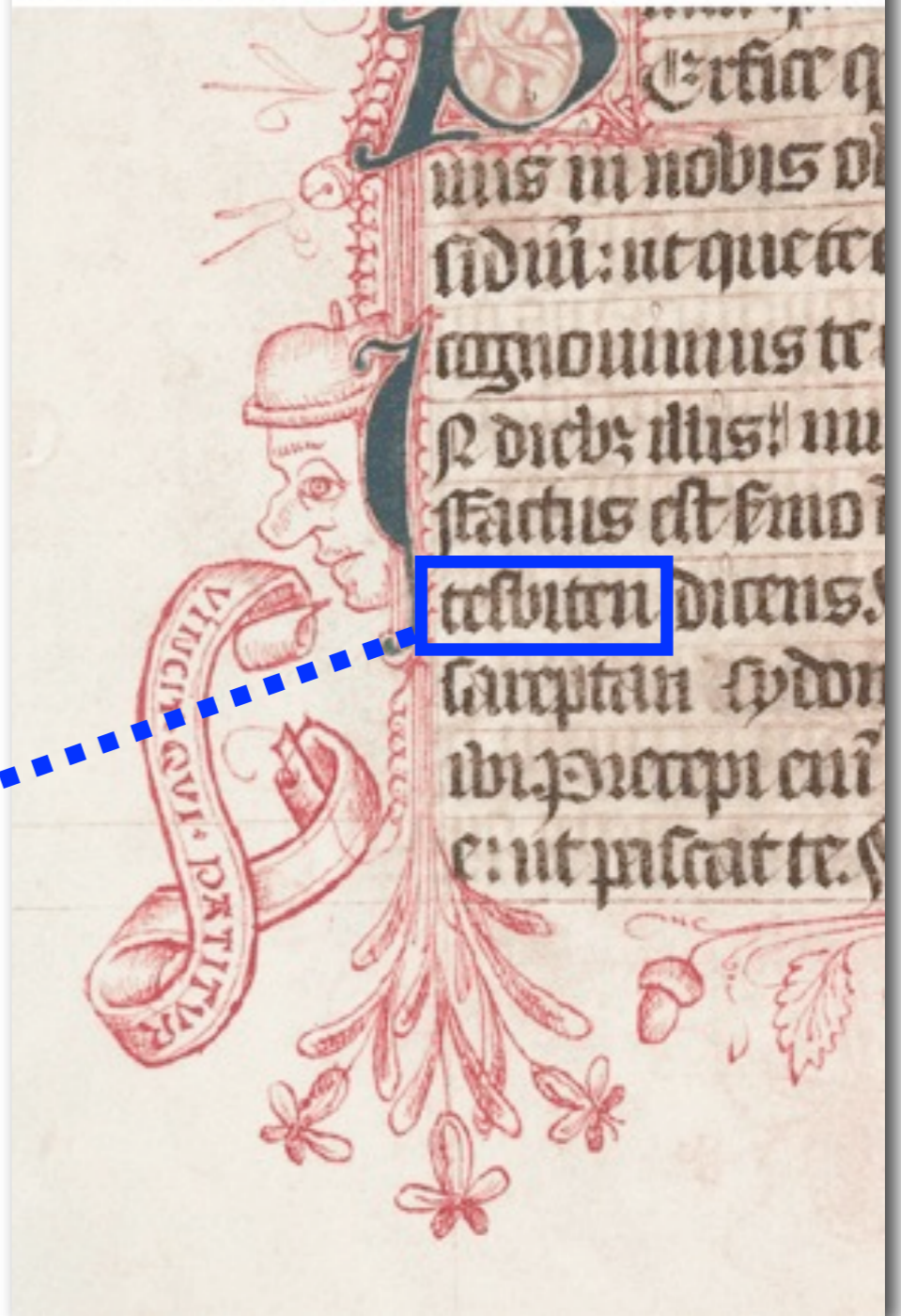
[Sign Up](#)

[Problems signing up? Check out our help pages](#)

Προβλήματα σημειώνοντας τα κωδικά ασφαλείας? Ελέγξτε τις σελίδες βοήθειας

[Σημειώστε](#)

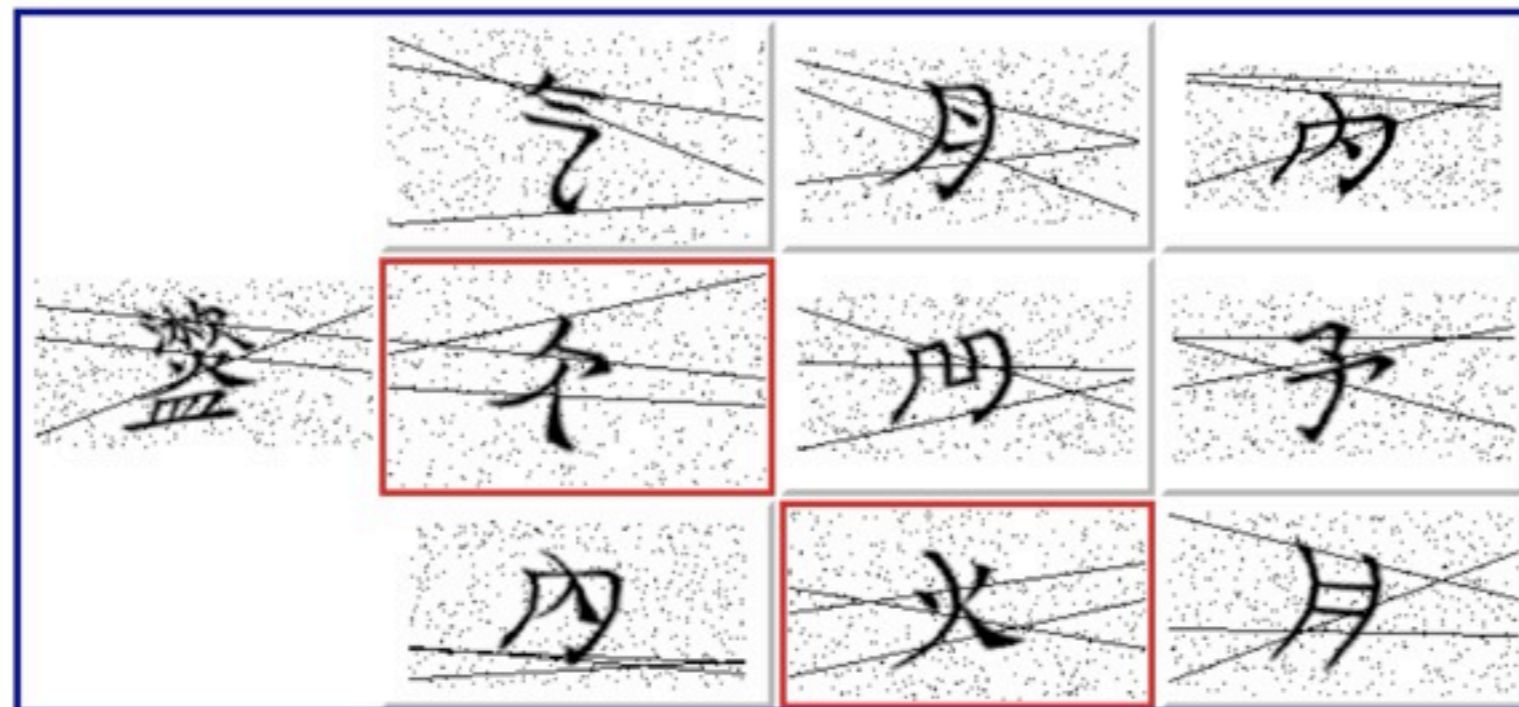
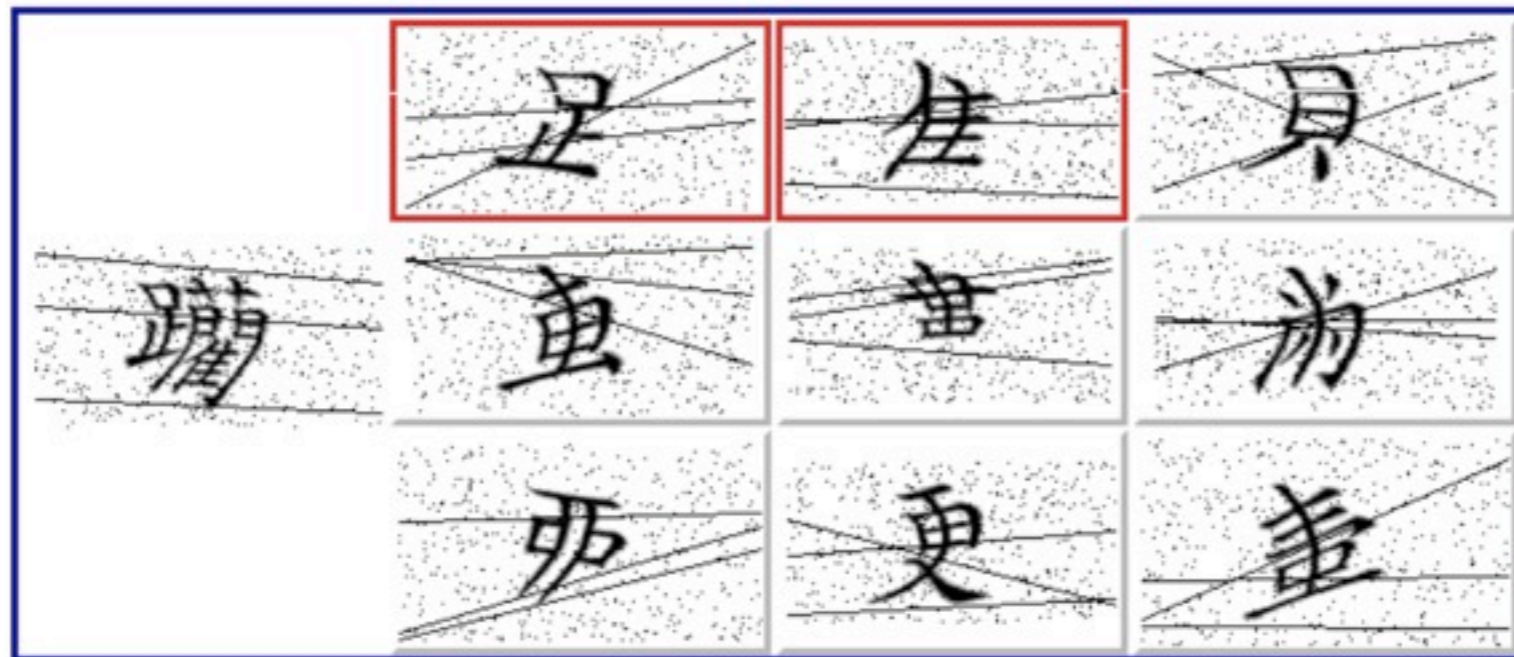
MS. Don. b. 6, fol. 48v (detail) © Bodleian Library, University of Oxford





# Chinese CAPTCHA

Ling-Jyh Chen, Institute of Information Science, Academia Sinica, Taipei, Taiwan



# Human Computation

New Game! How to Play GUEST

## PAGE HUNT

Cheetah CONSERVATION FUND

Search Partners | Links | Contact Us

Home About the Cheetah Who We Are What We Do How You Can Help News and Events Worldwide Locations

### HELP US SAVE THE WILD CHEETAH!

Our vision is to see a world in which cheetahs live and flourish in co-existence with people and the environment.

#### CCF'S MISSION

To be an internationally recognised centre of excellence in research cheetahs and their eco-systems, working with all stakeholders to a the conservation and management of the world's cheetahs.

Volunteer Kids4Cheetahs Cheetah

WHAT WILL IT TAKE TO SAVE THE CHEETAH? HELP TO CONSERVE

bing Copyright 2008 Microsoft Research | About Page Hunt | Terms of Use | Privacy Policy | Tell a Friend | Send Feedback

2:15 Score: 0 0 of 2 correct!

Frequent Queries: cheetah

cheetah fund bing

Query: cheetah fund Skip Bad Page

- 1 X [Cheetah Conservation Fund](#)  
The purpose of the Cheetah Conservation Fund (CCF) is to research and i  
<http://www.cheetah.org/?nd=home/>
- 2 X [<p>INTERNATIONAL LOCATIONS, PARTNERS, AFFILIAT](#)  
The purpose of the Cheetah Conservation Fund (CCF) is to research and i  
<http://www.cheetah.org/2nd-international/>



# Social Bookmarking

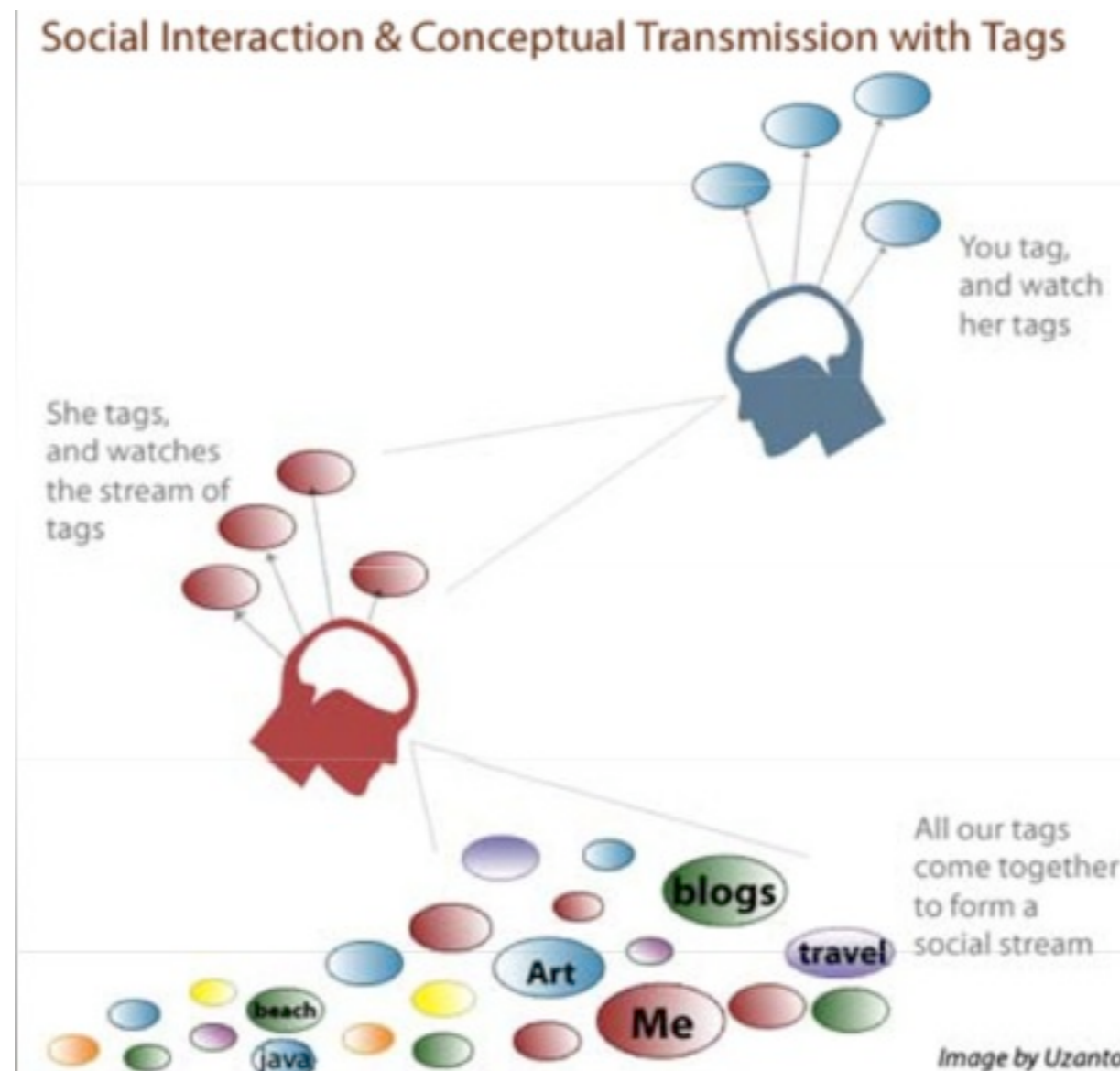
- **What is a tag?**
  - Descriptive metadata
  - A keyword or term associated with or assigned to a piece of information
  - User defined, created and shared
  - Many web users do it every day, with very little conscious awareness that they are “cataloging”
- **What gets tagged?**
  - Pictures, blog posts, video clips, catalog entries, just about anything...



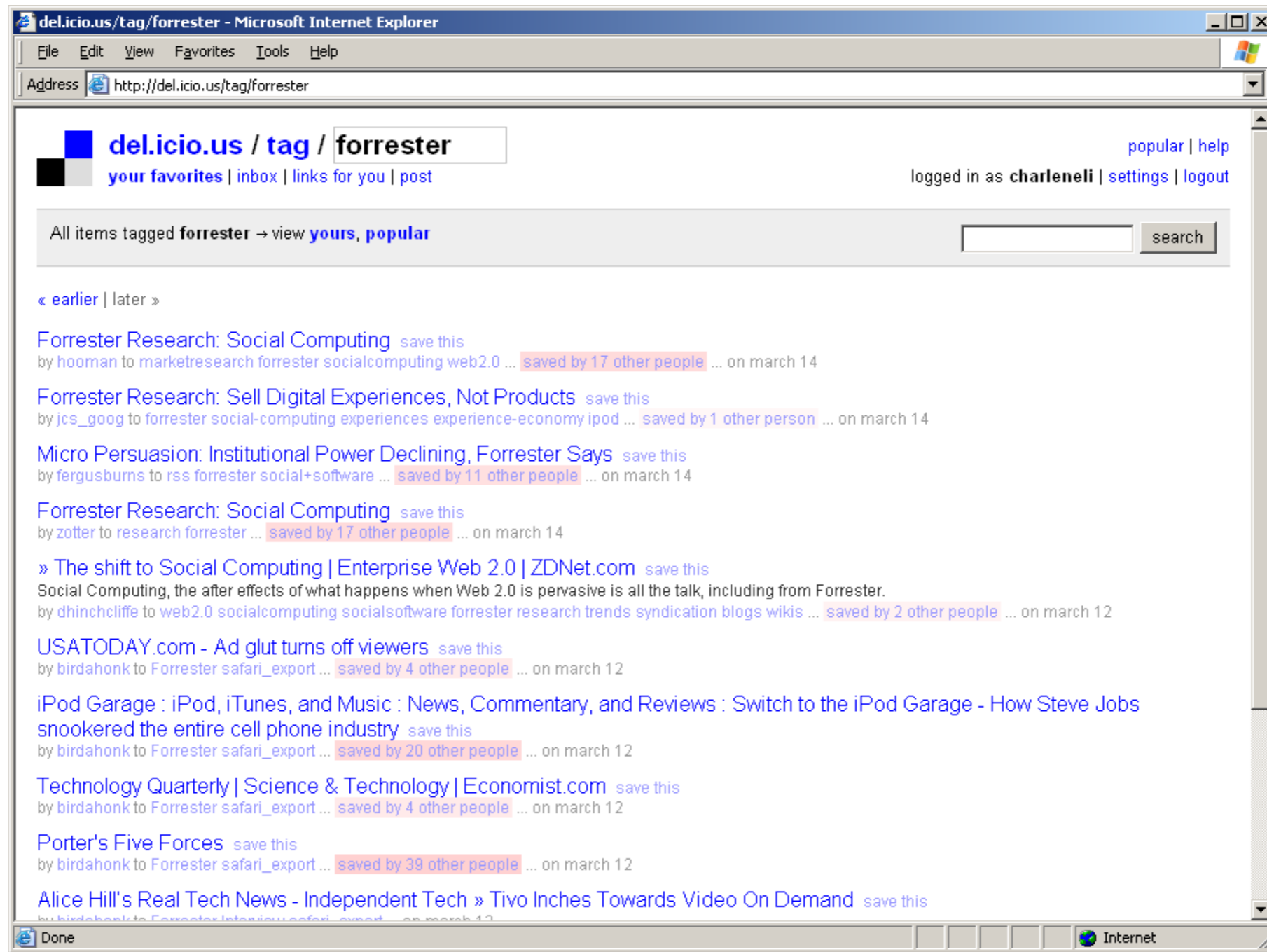


# Social Bookmarking

- Share one's tags
- Make the individual browsing experience a social one



# Social Bookmarking in del.icio.us





# Social Bookmarking in StumbleUpon

**StumbleUpon** allows users to discover and rate web pages, photos, and videos. It chooses which web page to display based on the user's ratings of previous pages, ratings by his/her friends, and by the ratings of users with similar interests.





# Tagging is Everywhere





# Social Entertainment

The screenshot shows the Swoopo website header with the logo, navigation links (Home, My Swoopo, Help, Register), and a login form. Below the header is a large banner for kitchenware. The banner features a kitchen scene with a stand mixer, knives, and vegetables. Text on the banner includes 'Starting NOW', 'CALPHALON, HENCKELS & KITCHENAID', and 'REGISTER NOW FOR FREE'. A 'Browse Kitchenware' link is also present.

The screenshot displays five auction items in a grid. Each item includes a title, a thumbnail image, a countdown timer, a current bid price, and a 'BID' button. The items are:

Item	Countdown	Current Bid	User
300 Bids Voucher	00:00:18	\$117.90	Nirajzala
MySims Agents (Nintendo DS)	00:02:05	\$0.24	Bb4kids
Samsung UN46B6000 46-Inch 1080p LED HDTV	00:00:15	\$102.00	Julia30
Wii   Nintendo Console + Wii Sports	00:00:15	\$32.04	Bearboy66
Apple MacBook Pro MB991LL/A 13.3-Inch Laptop	00:45:27	\$12.42	Jamesham



# Social Recommendations

## Genius Recommendations for Apps NEW

There are tens of thousands of apps in the App Store, with more added every day. A new feature of iPod touch makes finding cool new apps even easier. It's Genius for apps, and it works just like Genius for your music. Tap the Genius icon and get recommendations for apps that you might like based on apps you and others have downloaded.



## Genius Playlists

Say you're listening to a song you really like and want to hear other tracks that go great with it. The Genius feature finds other songs on your iPod touch that sound great with the one you were listening to and makes a Genius playlist for you. Listen to the playlist right away, save it for later, or even refresh it and give it another go. Count on Genius to create a mix you wouldn't have thought of yourself.

## Genius Mixes NEW

Now the Genius feature is even more powerful. Introducing Genius Mixes. All you do is sync iPod touch to iTunes, and Genius automatically searches your library to find songs that sound great together. Then it creates multiple mixes you'll love. These mixes are like channels programmed entirely with your music.





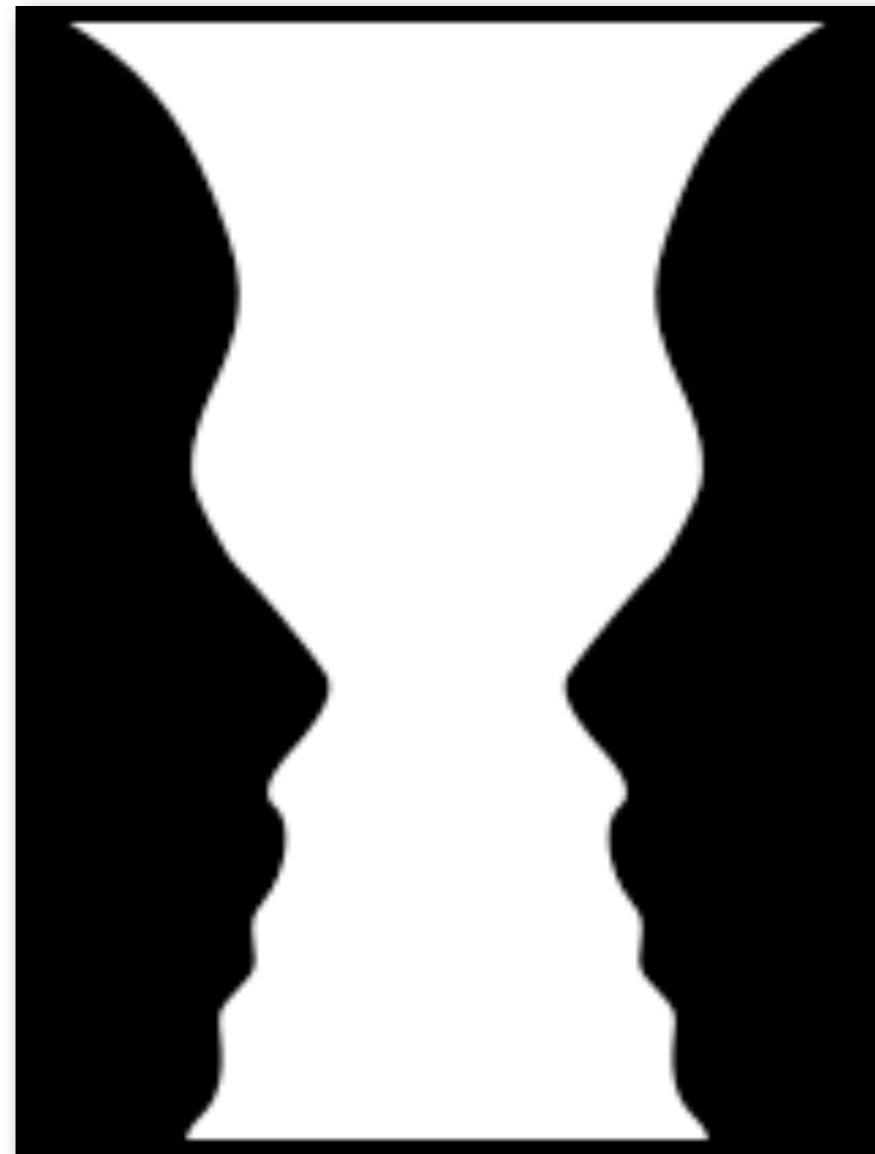
# Web 2.0 Revolution

- **Glocalization**-think globally and act locally!
- **Weblication**-Web is the application!
- Three C's

**C**onnectivity

**C**ollaboration

**C**ommunities



# Social Relations

presence  
identity  
social role  
reputation  
expertise  
trust  
ownership  
accountability  
knowledge

crew  
teams  
squad  
cohorts  
communities  
groups

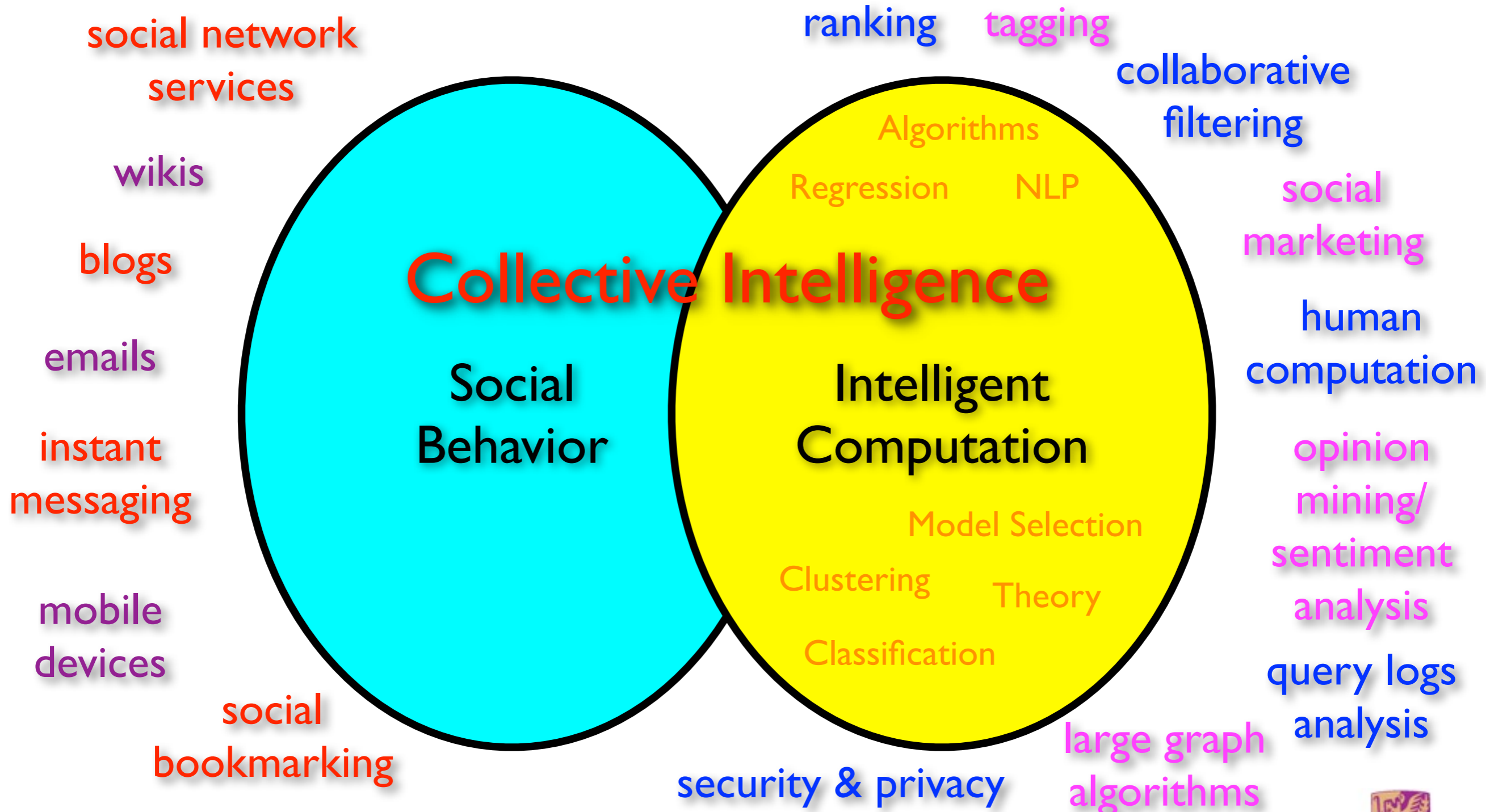
populations  
organizations  
markets  
partners

binary  
cardinal  
integer  
real





# Social Recommendation



# Emerging Issues

- **Theory** and models
- **Search, mining, ranking and recommending** of existing information, e.g., spatial (relations) and temporal (time) domains
  - Dealing with **partial** and **incomplete** information, e.g., collaborative filtering, ranking, tagging, etc.
- **Scalability** and algorithmic issues
- **Security** and **privacy** issues
- **Monetization** of social interactions





# Introduction

- Social Platforms
- Techniques in Social Recommendation
  - Social Network Theory
  - Graph/Link Mining
  - Collaborative Filtering
  - Machine Learning Techniques
- Summary



# Social Network Theory

- Consider many kinds of networks:
  - social, technological, business, economic, content, ...
- These networks tend to share certain informal properties:
  - **large scale**; continual growth
  - **distributed**, organic growth: vertices “decide” who to link to
  - interaction restricted to **links**
  - mixture of **local** and **long-distance** connections
  - **abstract** notions of distance: geographical, content, social,...





# Social Network Theory

- Do these networks share more **quantitative** universals?
- What would these “universals” be?
- How can we make them precise and measure them?
- How can we explain their universality?
- This is the domain of **social network theory**



# Some Interesting Quantities

- **Connected components**
  - how many, and how large?
- **Network diameter**
  - maximum (worst-case) or average?
  - exclude infinite distances? (disconnected components)
  - the small-world phenomenon
- **Clustering**
  - to what extent that links tend to cluster “locally”?
  - what is the balance between local and long-distance connections?
  - what roles do the two types of links play?
- **Degree distribution**
  - what is the typical degree in the network?
  - what is the overall distribution?





# Graph/Link Mining

- Heterogeneous, multi-relational data represented as a graph or network
  - Nodes are objects
    - Objects have attributes
    - Objects may have labels or classes
  - Edges are links
    - Links may have attributes
    - Links may be directed
- Links represent relationships and interactions between objects -- rich content for mining



# What Is New For Mining

- Traditional machine learning and data mining approaches assume:
  - A random sample of homogeneous objects from single relation
- Real world data sets:
  - **Multi-relational, heterogeneous and semi-structured**
- Link Mining
  - Newly emerging research area at the intersection of research in social network and link analysis, hypertext and web mining, graph mining, relational learning and inductive logic programming



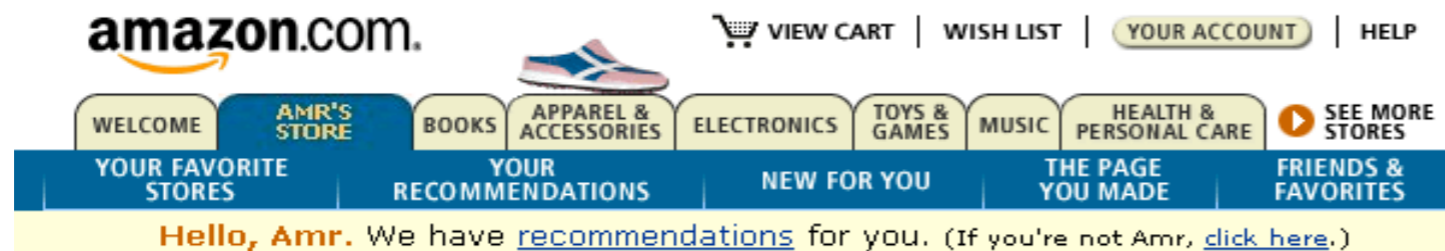


# What is a Link in Link Mining

- Link: relationship among data
- **Homogeneous networks**
  - Single object type and single link type
  - Single model social networks (e.g., friends)
  - WWW: a collection of linked Web pages
- **Heterogeneous networks**
  - Multiple object and link types
  - Medical network: patients, doctors, disease, contacts, treatments
  - Bibliographic network: publications, authors, venues



# Real life Example for Collaborative Filtering



## [The Page You Made](#)



[Understanding Search Engines](#)  
by Michael W. Berry, Murray Browne  
Price: **\$41.50**

Book News, Inc.

Berry and Browne (computer science, U. of Tennessee) discuss key design issues in information retrieval about which their computer science peers and... [Read more](#) | [Why was I recommended](#)

- User's perspective
  - **Lots** of online products, books, movies, etc
  - **Reduce** my choices

- Manager's perspective

*"if I have 3 million **customers** on the web, I should have 3 million **stores** on the web."*

CEO of Amazon.com



# More Examples

- **Movielens**: movies
- **Moviecritic**: movies again
- **My launch**: music
- **Gustos starrater**: web pages
- **Jester**: Jokes
- **TV Recommender**: TV shows
- **Suggest 1.0**: different products
- And much more...





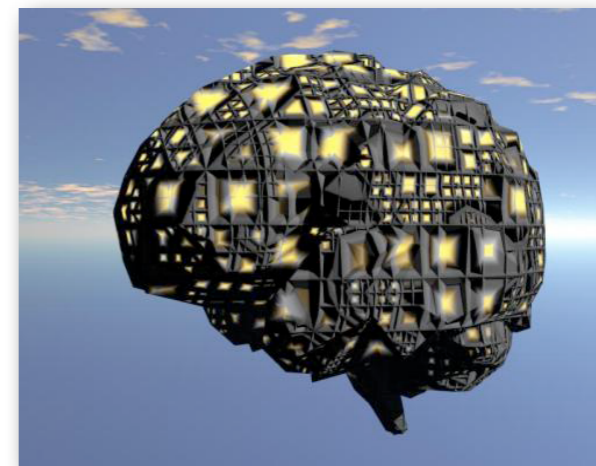
# How it Works?

- Each user has a **profile**
- Users **rate** items
  - Explicitly: score from 1..5
  - Implicitly: web usage mining
    - **Time** spent in viewing the item
    - Navigation path, etc...
- System does the rest, How?
  - Look at users **collective** behavior
  - Look at the active user **history**

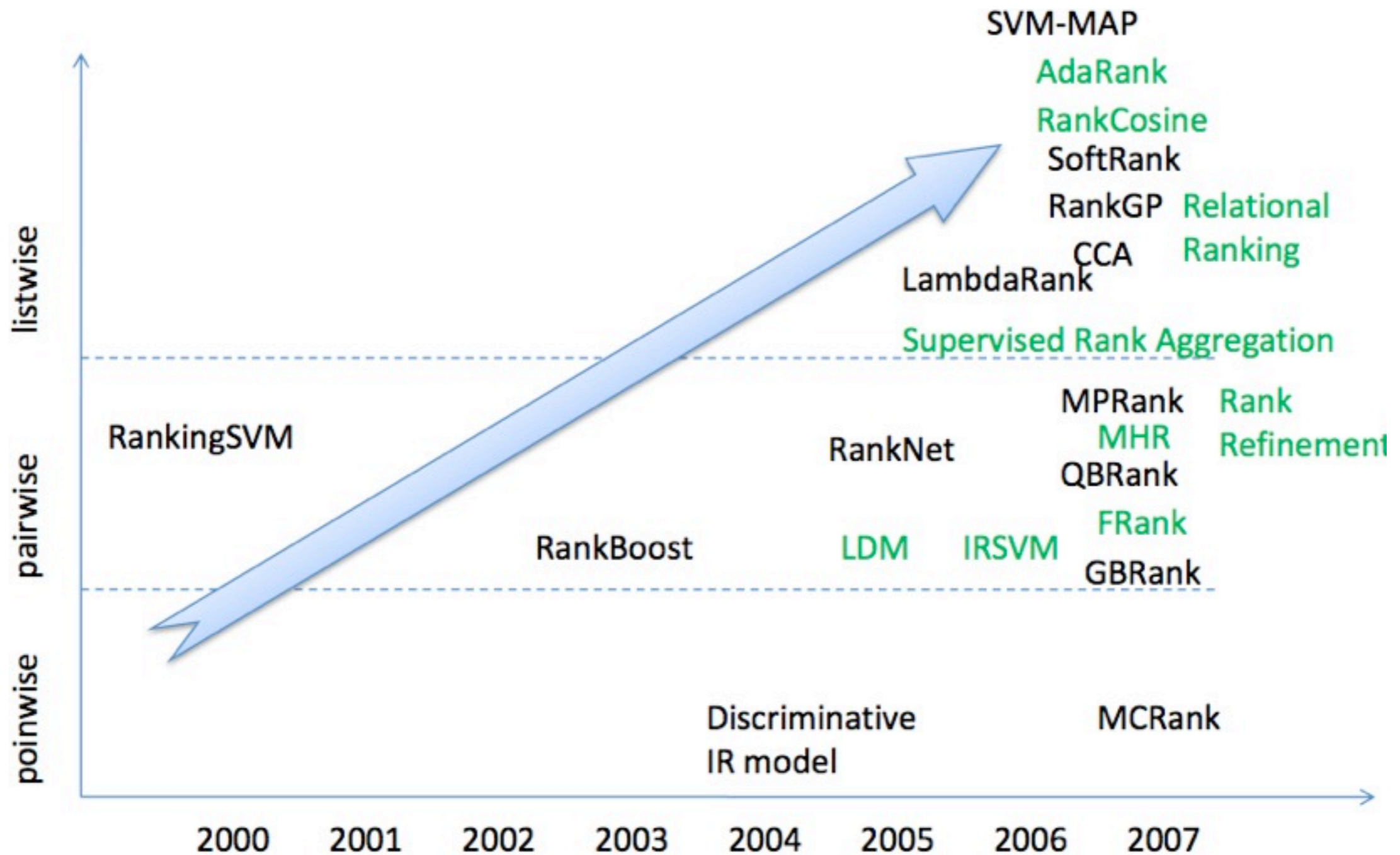


# Machine Learning Can Help

- Machine learning is an effective tool
  - To automatically tune parameters
  - To combine multiple evidences
  - To avoid over-fitting (by means of regularization, etc.)
- **Learning to Rank**
  - Use machine learning technologies to train the ranking model
  - A hot research topic these years



# Learning To Rank Techniques



<http://research.microsoft.com/en-us/people/tyliu/default.aspx>

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA





# Summary

- **Social Platforms**
  - Social Network
  - Social Media
  - Social games
  - Social bookmarking
  - Social News and Social Knowledge Sharing
- **Techniques in Social Recommendation**
  - Social Network Theory
  - Graph/Link Mining
  - Collaborative Filtering
  - Machine Learning Techniques



# References

- <https://agora.cs.illinois.edu/display/cs512/home>.
- 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- 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.
- R. Jaschke, M. Grahl, A. Hotho, B. Krause, C. Schmitz, and G. Stumme. Organizing publications and bookmarks in bibsonomy. In CKC, 2007.
- L. von Ahn. Games with a purpose. IEEE Computer, 39(6):92–94, 2006.



# References

- C. S. Andreas Hotho, Robert Jaschke and G. Stumme I. Bibsonomy: A social bookmark and publication sharing system. In CS-TIW'06. Aalborg University Press, 2006.
- G. W. Furnas, C. Fake, L. von Ahn, J. Schachter, S. A. Golder, K. Fox, M. Davis, C. Marlow, and M. Naaman. Why do tagging systems work? In CHI Extended Abstracts, pages 36–39, 2006.
- P. Heymann, G. Koutrika, and H. Garcia-Molina. Can social bookmarking improve web search? In WSDM, pages 195–206, 2008.
- R. Jaschke, L. B. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag recommendations in folksonomies. In A. Hinneburg, editor, LWA, pages 13–20. 2007.
- B. Krause, A. Hotho, and G. Stumme. A comparison of social bookmarking with traditional search. In ECIR, pages 101–113, 2008.
- L. Specia and E. Motta. Integrating folksonomies with the semantic web. In ESWC, pages 624–639, 2007.





# References

- R. B. D. M. C. Edith L. M. Law, Luis von Ahn. Tagatune: A game for music and sound annotation. ISMIP, 2007.
- L. von Ahn and L. Dabbish. Labeling images with a computer game. In CHI, pages 319–326, 2004.
- L. von Ahn and L. Dabbish. Designing games with a purpose. Commun.ACM, 51(8): 58–67, 2008.
- L. von Ahn, S. Ginosar, M. Kedia, R. Liu, and M. Blum. Improving accessibility of the web with a computer game. In CHI, pages 79–82, 2006.
- L. von Ahn, M. Kedia, and M. Blum. Verbosity: a game for collecting common-sense facts. In CHI, pages 75–78, 2006.
- L. von Ahn, R. Liu, and M. Blum. Peekaboom: a game for locating objects in images. In CHI '06, pages 55–64, New York, NY, USA, 2006. ACM.
- H. Ma, R. Chandrasekar, C. Quirk, A. Gupta: Improving search engines using human computation games. In CIKM '09, pages 275–284, Hong Kong, 2009



# References

- Ashwin Machanavajjhala , Daniel Kifer , Johannes Gehrke , Muthuramakrishnan Venkatasubramanian, L-diversity: Privacy beyond k-anonymity, TKDD, 2007
- Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian, t-Closeness: Privacy Beyond k-Anonymity and l-Diversity, ICDE, 2007.
- Xiao, X., Tao, Y, Dynamic Anonymization: Accurate Statistical Analysis with Privacy Preservation, SIGMOD, 2008.
- Michael Hay, Gerome Miklau, David Jensen, Don Towsley and Philipp Weis, Resisting Structural Re-identification in Anonymized Social Networks, PVLDB, 2008
- Lars Backstrom, Cynthia Dwork and Jon Kleinberg, Wherefore Art Thou R3579X? Anonymized Social Networks, Hidden Patterns, and Structural Steganography, WWW, 2007
- Kun liu and Evimaria Terzi, Towards Identity Anonymization on Graphs. SIGMOD, 2008
- Bin Zhou and Jian Pei, Preserving Privacy in Social Networks Against Neighborhood Attacks, ICDE, 2008



# Outline

- Introduction
- **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/](http://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](http://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/](http://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** resurrection 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





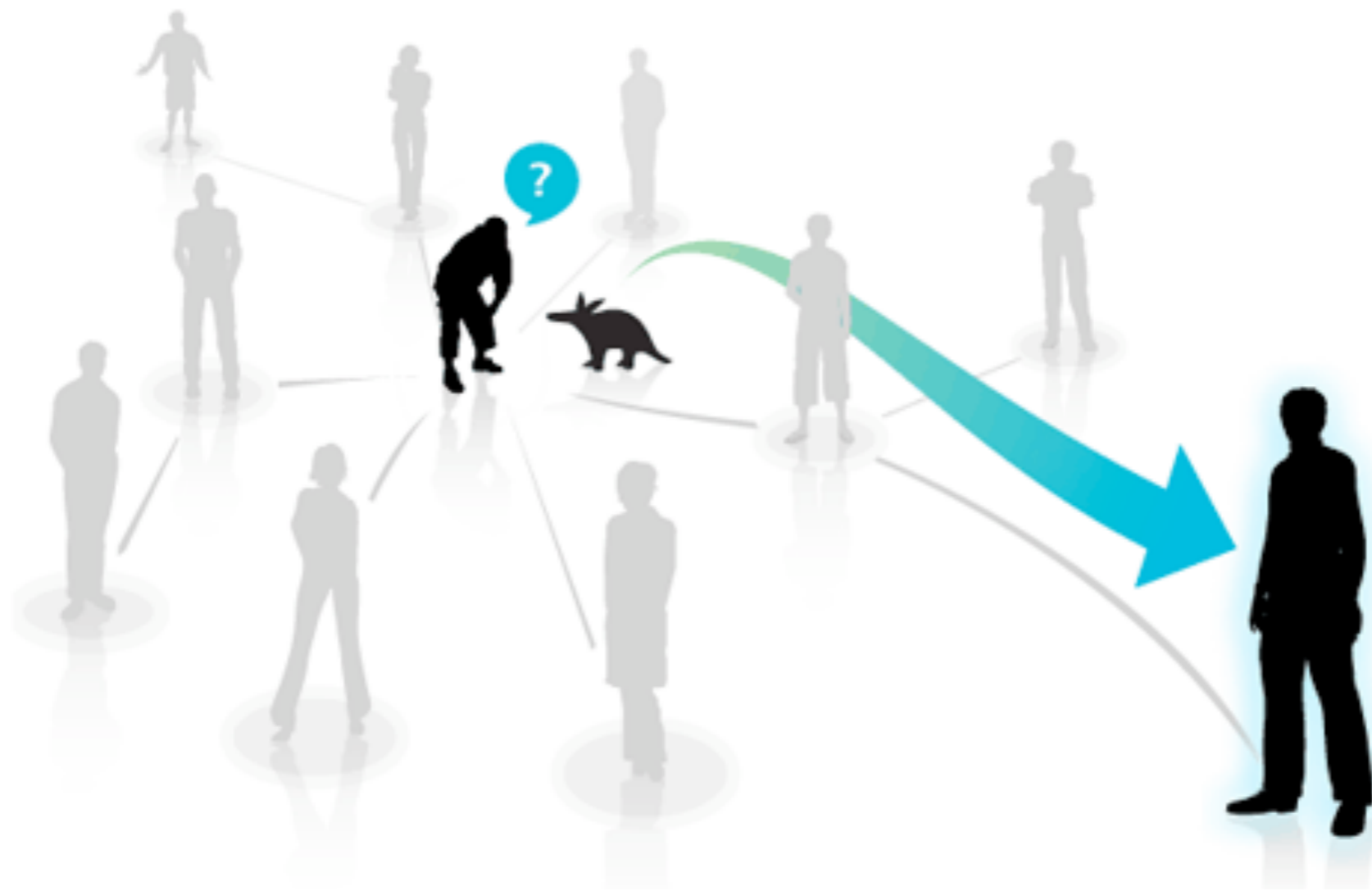
# Evolution of Search

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



# The Anatomy of A Large-Scale Social Search Engine

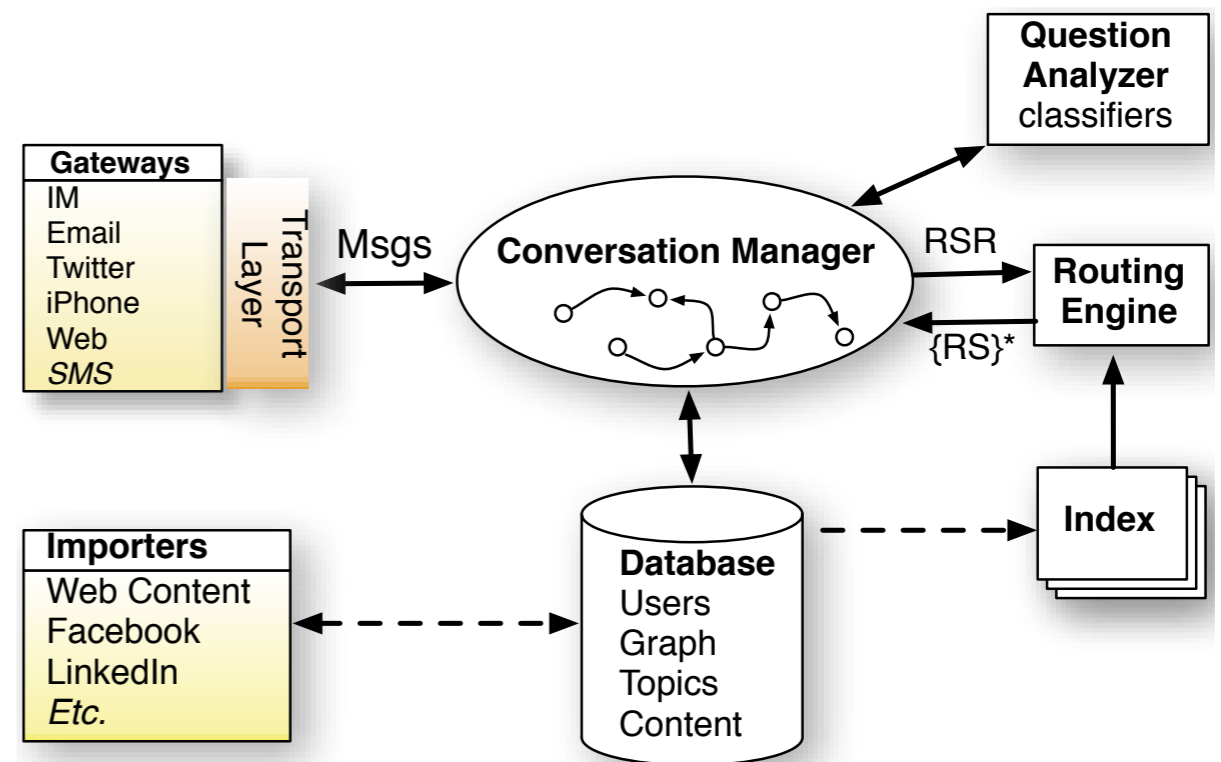
[D. Horowitz et al., WWW2010]



# 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

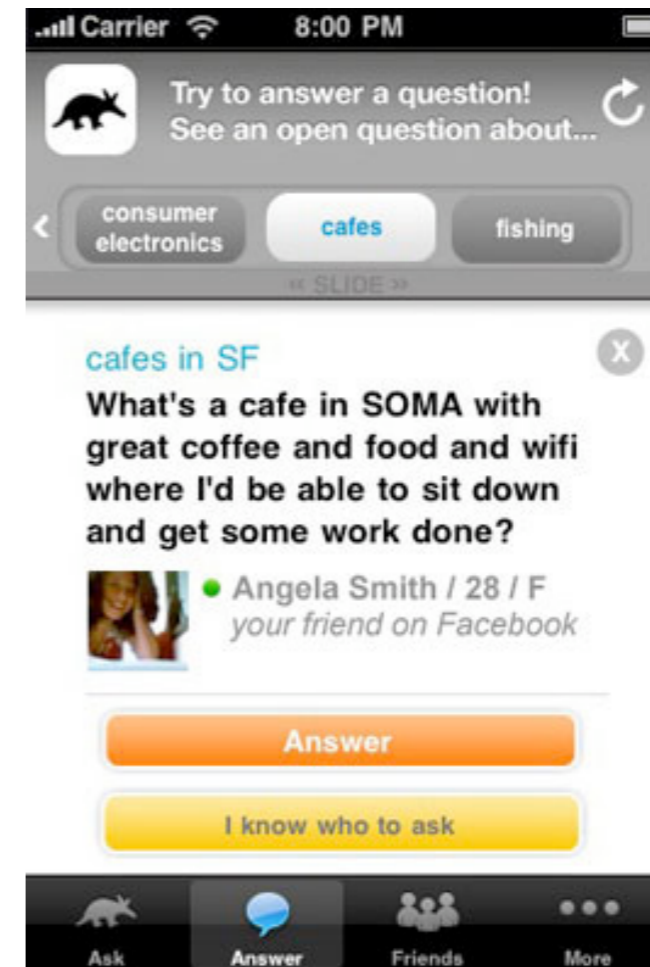


Figure 4: Screenshot of Aardvark Answering Tab on iPhone



# The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]

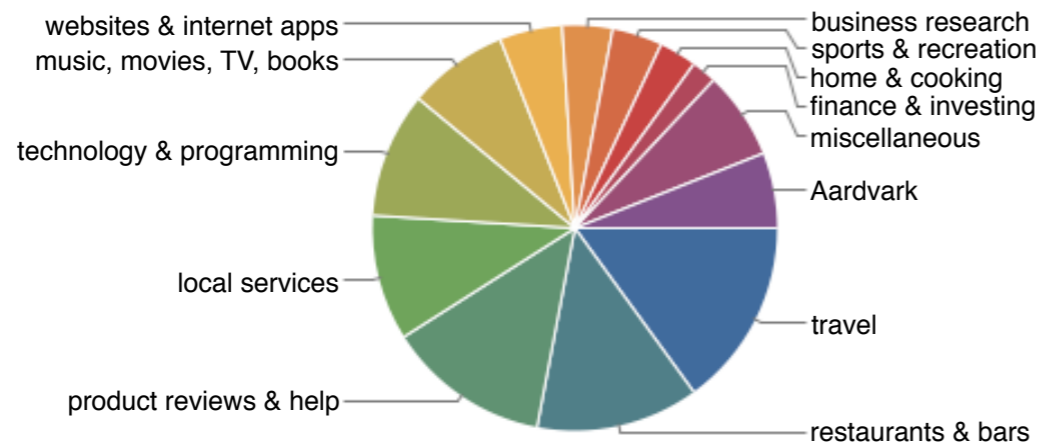


Figure 8: Categories of questions sent to Aardvark

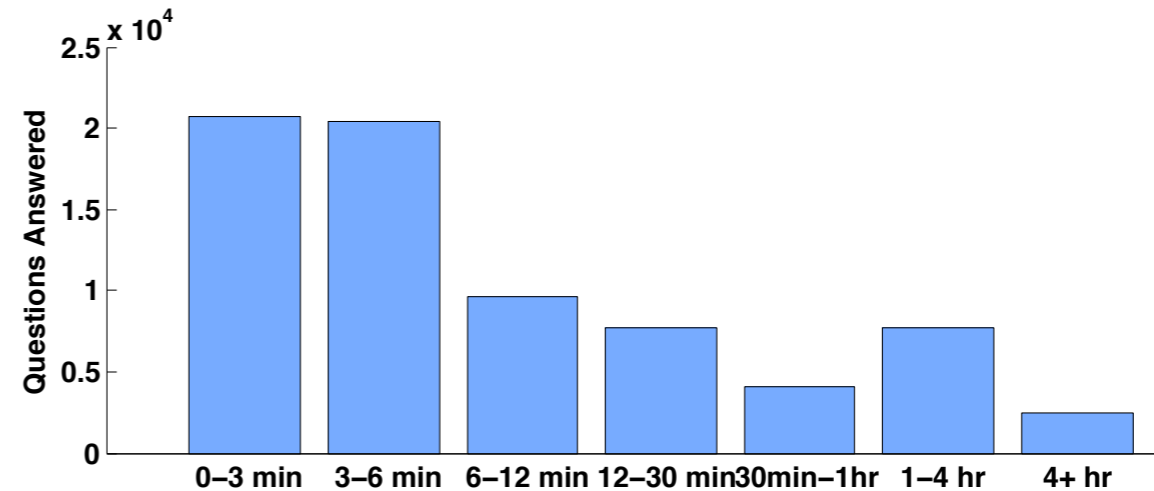


Figure 9: Distribution of questions and answering times.

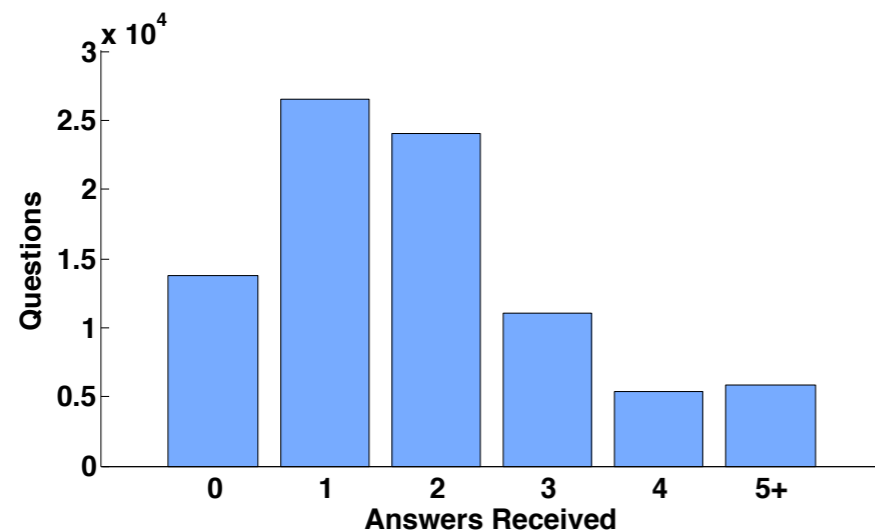


Figure 10: Distribution of questions and number of answers received.

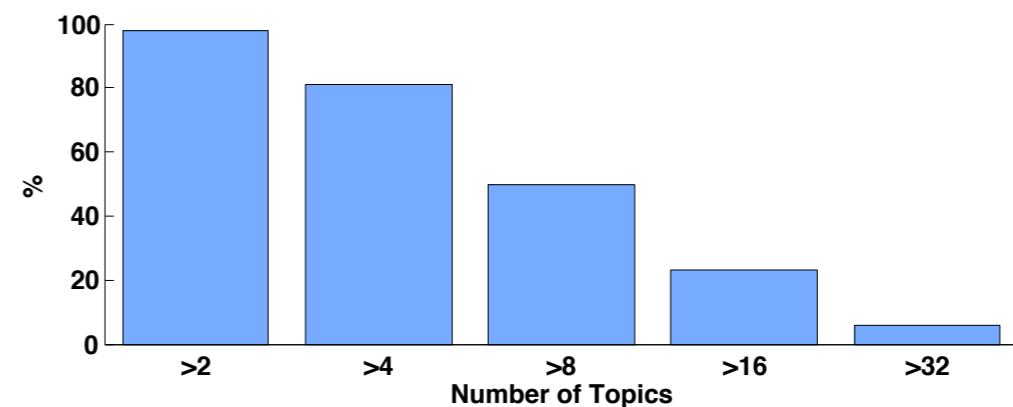


Figure 11: Distribution of percentage of users and number of topics





# 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

- Introduction
- 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



Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA





# Real Life Examples

The screenshot shows the Amazon.com product page for the book "Social Computing, Behavioral Modeling, and Prediction (Hardcover)". The page includes a "Click to LOOK INSIDE!" button with a magnifying glass icon over the book cover. The book is edited by Huan Liu, John J. Salerno, and Michael J. Young. Key phrases include "social network analysis", "electronic institutions", and "cognitive modeling". The list price is \$129.00, and the current price is \$129.00 with free shipping. The book is in stock and ships from Amazon.com. A "Frequently Bought Together" section is circled in red, showing the book being purchased with "Social Network Analysis: A Handbook" by John P. Scott for a total price of \$174.85. The right sidebar contains purchase options, including "Add to Shopping Cart", "Add to Cart with FREE Two-Day Shipping", and "More Buying Choices" (16 used & new from \$95.29). Other options include "Add to Wish List", "Add to Shopping List", "Add to Wedding Registry", "Add to Baby Registry", and "Share with Friends".

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.](#)

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

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](#)





# 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** by Ariely

**Predictably Irrational: Hidden Forces That...** by Ariely

★★★★☆ (18) \$16.34

LOOK INSIDE! **COMPLEX**

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

★★★★☆ (5) \$42.00

**Editorial Reviews**

**Product Description**


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

Five scales rating

- ★ I hate it
- ★★ I don't like it
- ★★★ It's ok
- ★★★★ I like it
- ★★★★★ I love it



# 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 Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA



# 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



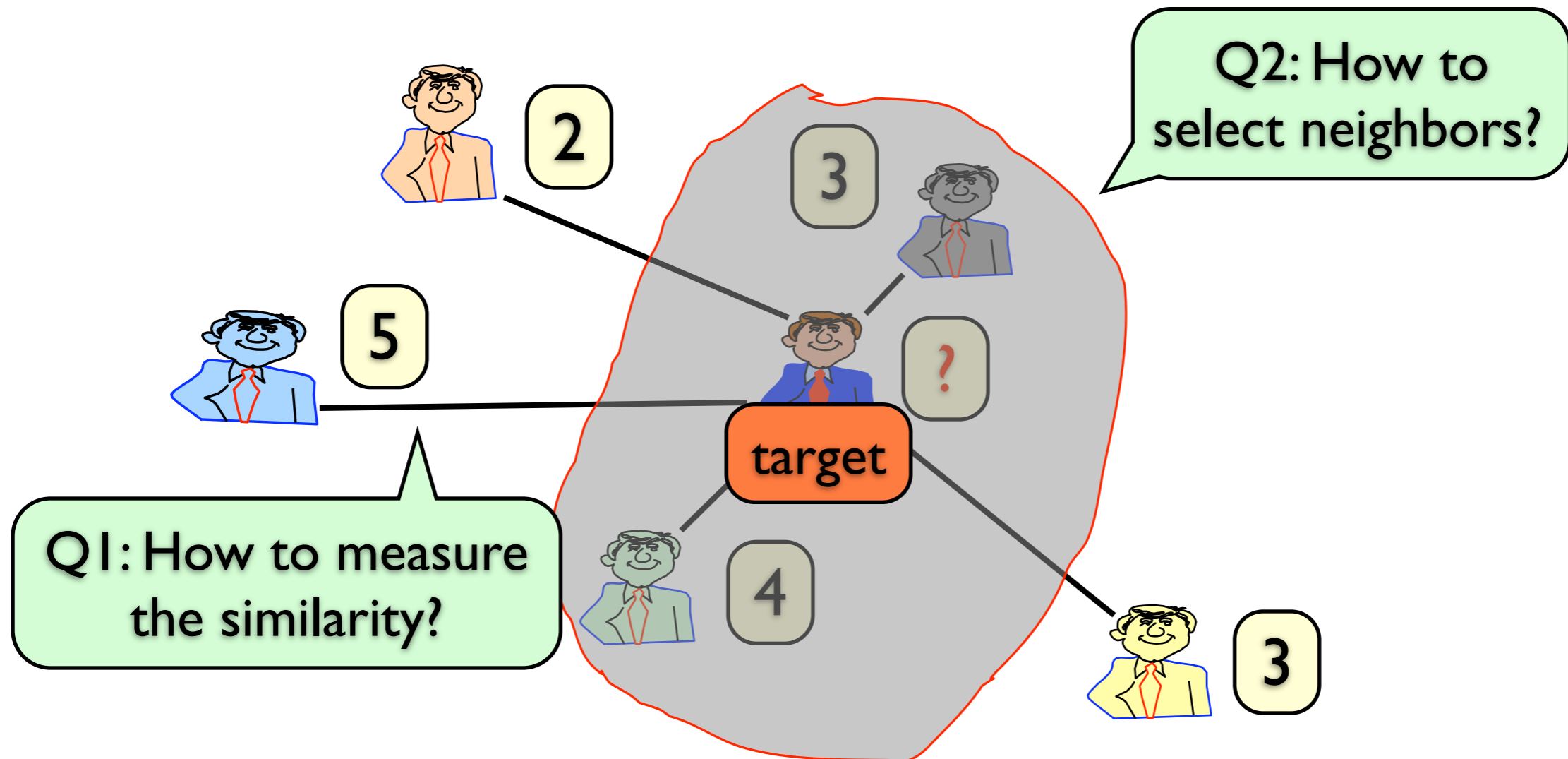
# Collaborative Filtering

- 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 <sub>1</sub>													
u <sub>2</sub>	1	3		4		2		5			3	4	
u <sub>3</sub>													
u <sub>4</sub>		3		4			3	4		3	4		4
u <sub>5</sub>													
u <sub>6</sub>	1			3	5	2		4	1			3	



# User-based Collaborative Filtering

Items

Users

u <sub>1</sub>												
u <sub>2</sub>	1	3	4		2		5			3	4	
u <sub>3</sub>												
u <sub>4</sub>		3	4			3	4		3	4		4
u <sub>5</sub>												
u <sub>6</sub>	1		3	5	2		4	1			3	





# User-based Collaborative Filtering

Items

Users

u <sub>1</sub>													
u <sub>2</sub>	1	3	4	2	5			3	4				
u <sub>3</sub>													
u <sub>4</sub>		3	4		3	4		3	4		4		
u <sub>5</sub>													
u <sub>6</sub>	1		3	5	2	4	1				3		



# User-based Collaborative Filtering

Items

Users

u <sub>1</sub>													
u <sub>2</sub>	1	3		4		2		5			3	4	
u <sub>3</sub>													
u <sub>4</sub>		3		4			3	4		3	4		4
u <sub>5</sub>													
u <sub>6</sub>	1			3	5	2		4	1			3	



# User-based Collaborative Filtering

Items

Users

u <sub>1</sub>													
u <sub>2</sub>	1	3	4		2		5			3	4		
u <sub>3</sub>													
u <sub>4</sub>		3	4			3	4		3	4		4	
u <sub>5</sub>													
u <sub>6</sub>	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	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  $\lambda_1, \lambda_2 > 0$ .

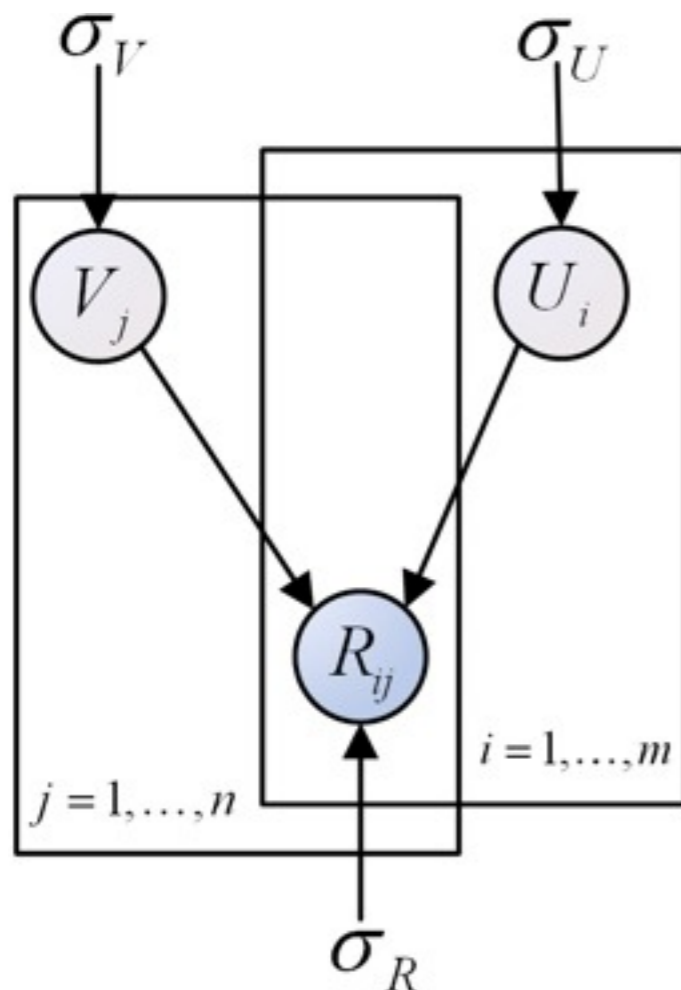
- The problem can be solved by simple gradient descent algorithm.





# Probabilistic Matrix Factorization

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

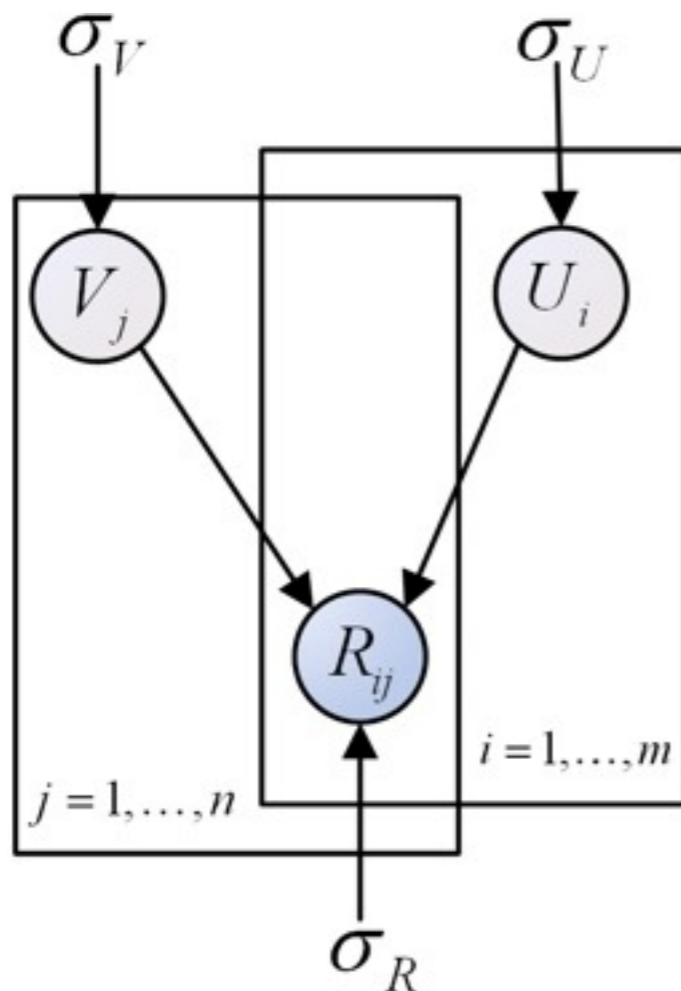


$$p(R|U, V, \sigma_R^2) = \prod_{i=1}^m \prod_{j=1}^n \left[ \mathcal{N} \left( R_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right]^{I_{ij}^R}$$



# Probabilistic Matrix Factorization

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



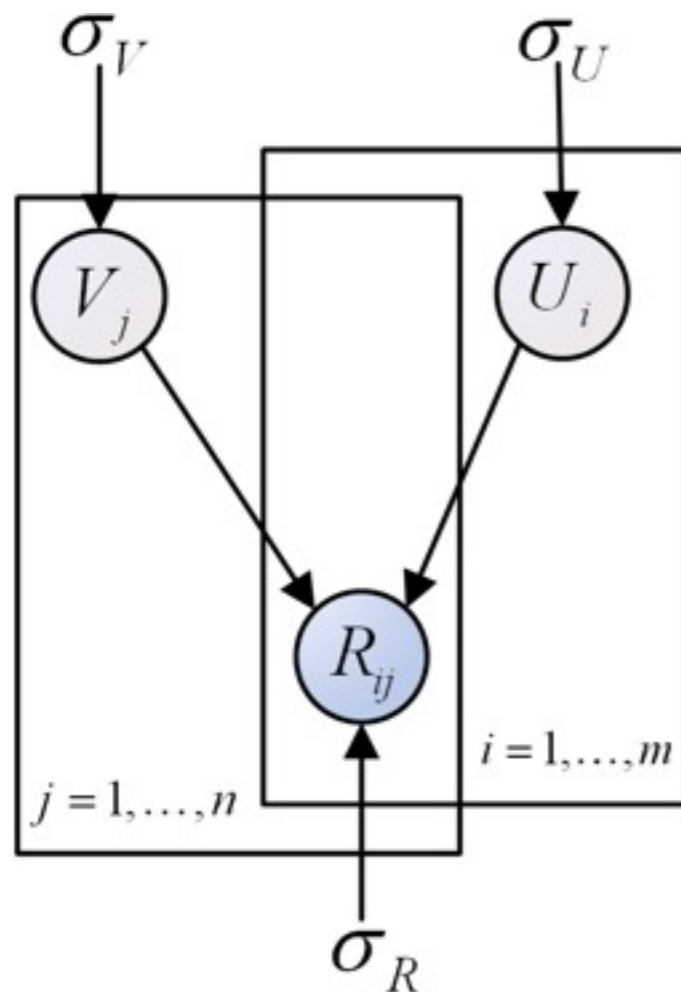
$$p(U|\sigma_U^2) = \prod_{i=1}^m \mathcal{N}(U_i|0, \sigma_U^2 \mathbf{I})$$

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



# 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



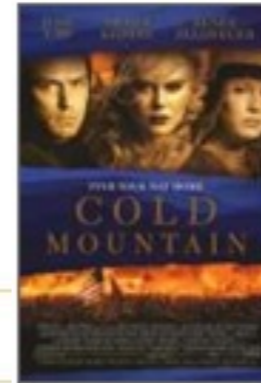
**My Blueberry Nights** (PG-13, 1 hr. 30 min.)

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

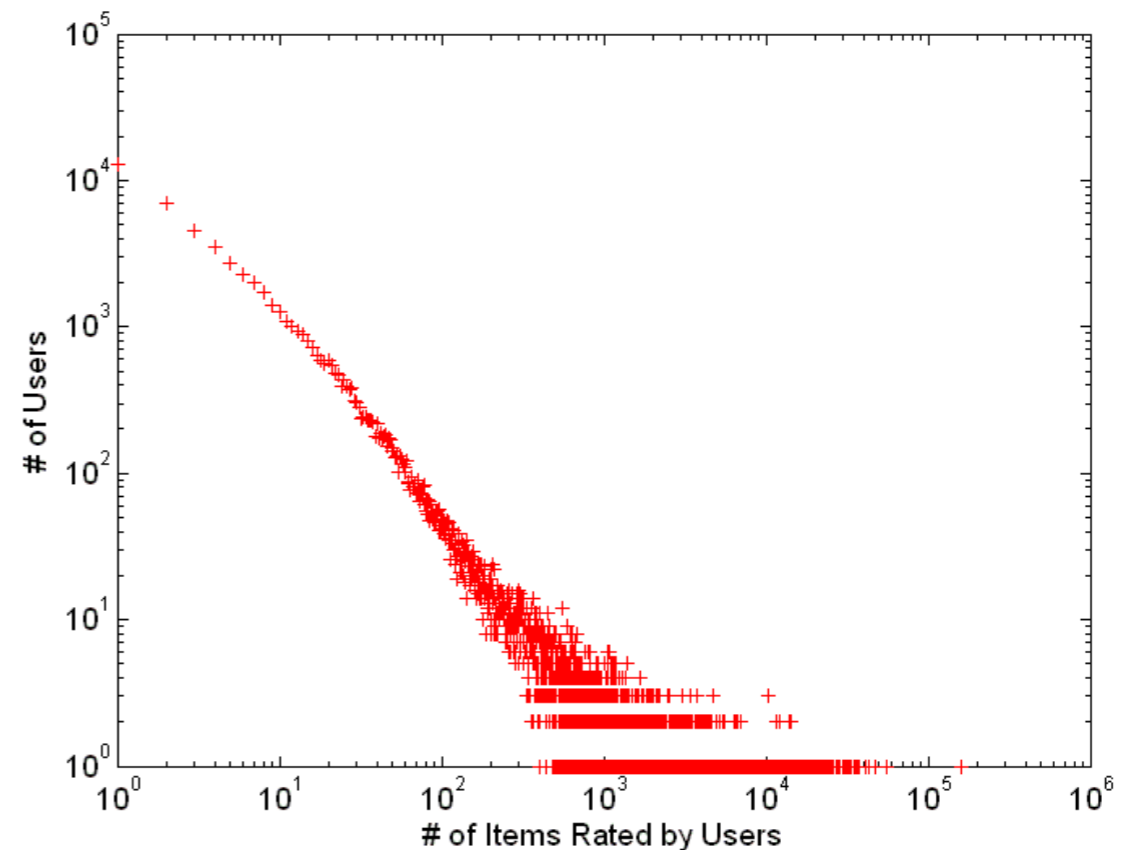
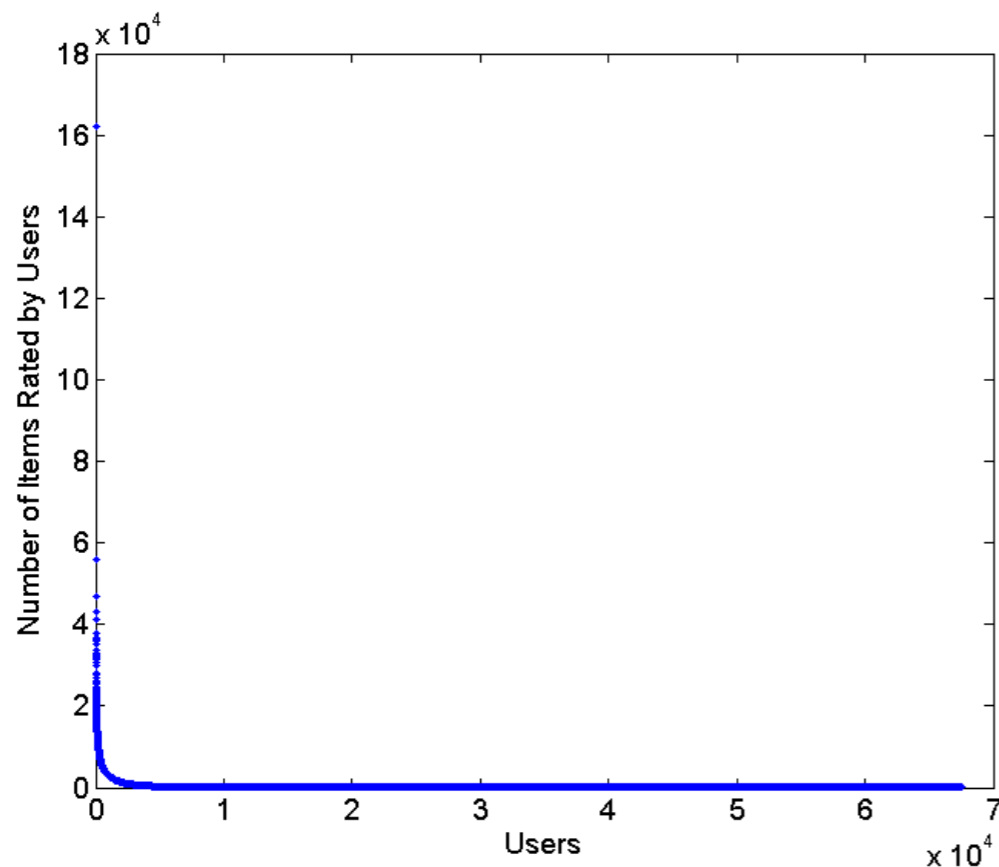
The Critics: **B** 15 reviews

★ My Rating: B





# Number of Ratings per User



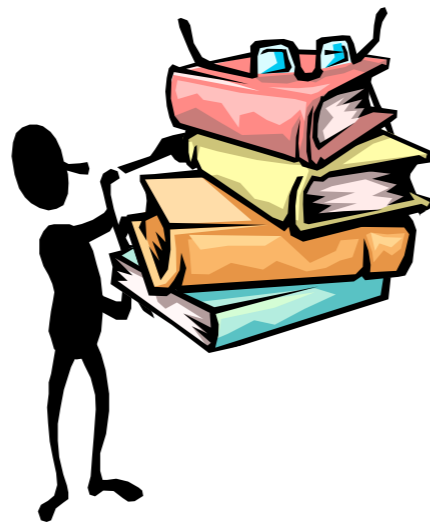
Extracted From [Epinions.com](http://Epinions.com)  
114,222 users, 754,987 items and 13,385,713 ratings





# Challenges

- Traditional recommender systems ignore the social connections between users



Which one should I choose?

Recommendations  
from friends



# Social Recommendation Using Probabilistic Matrix Factorization

[Hao Ma, et al., CIKM2008]



# Motivations

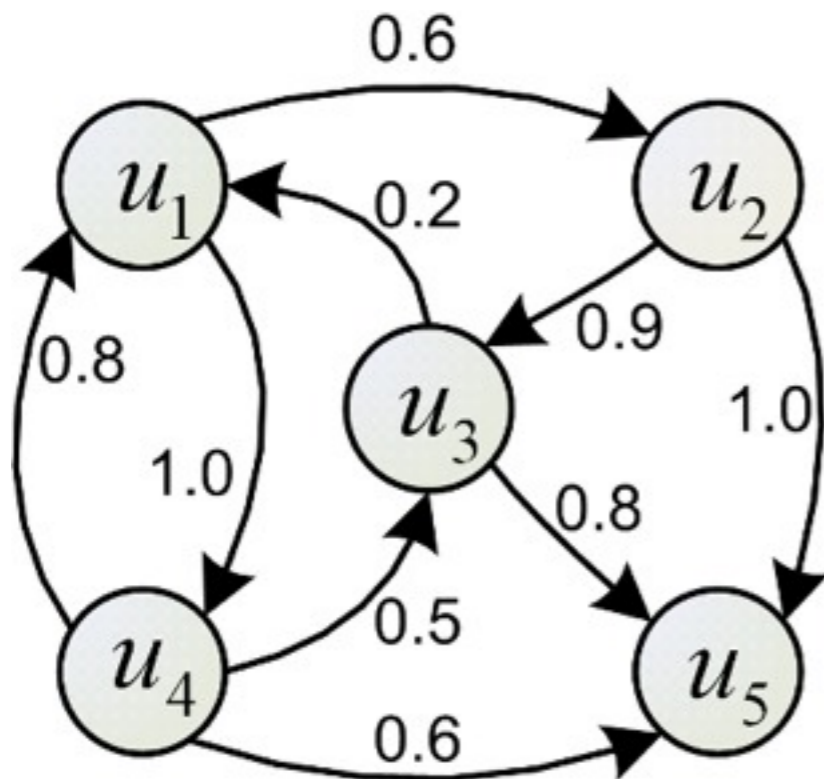
- “Yes, there is a correlation - from social networks to personal behavior on the web”

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

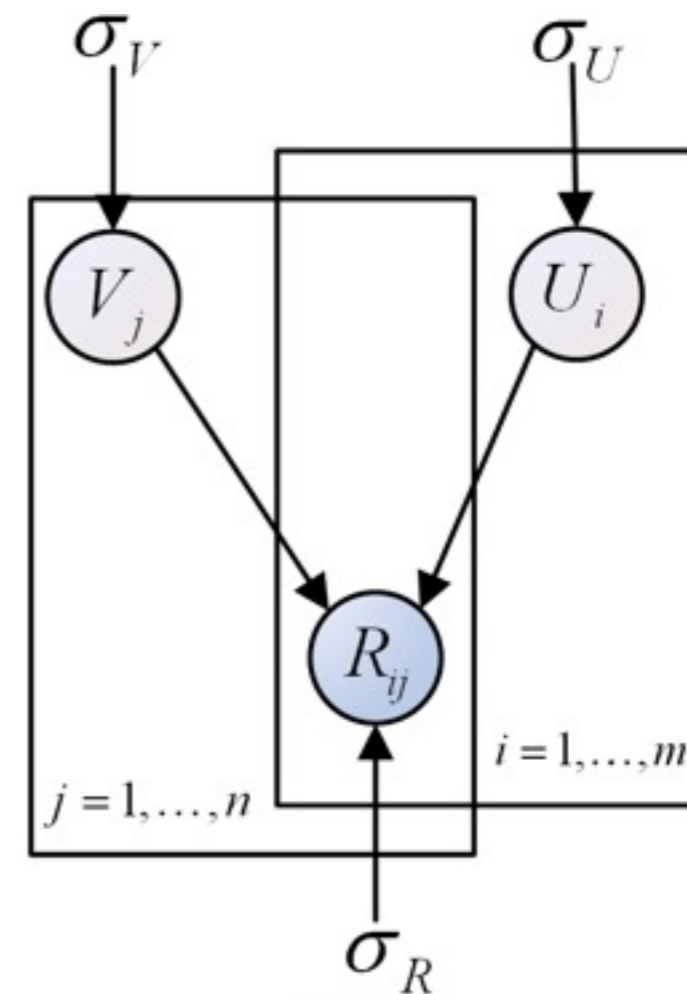
User-Item Rating Matrix





# 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



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

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

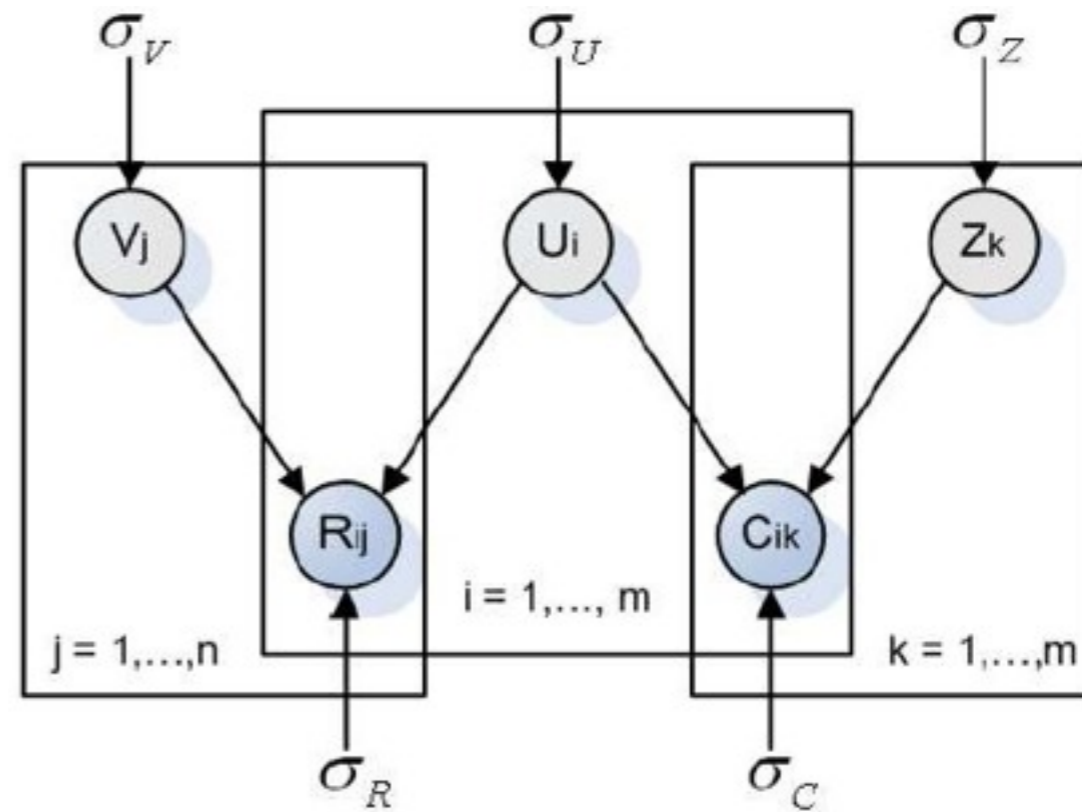
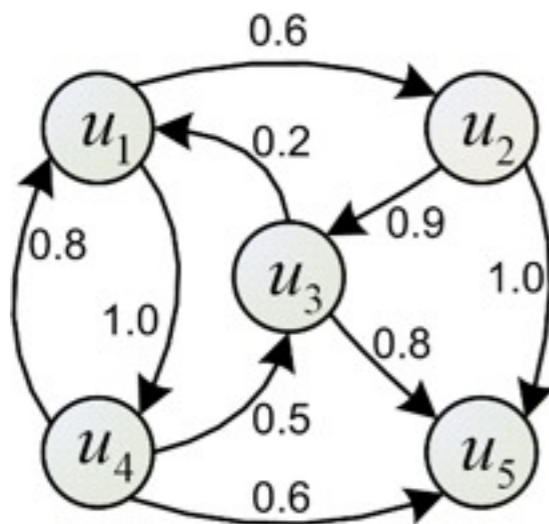
$$p(V|\sigma_V^2) = \prod_{j=1}^n \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I})$$

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



# SoRec

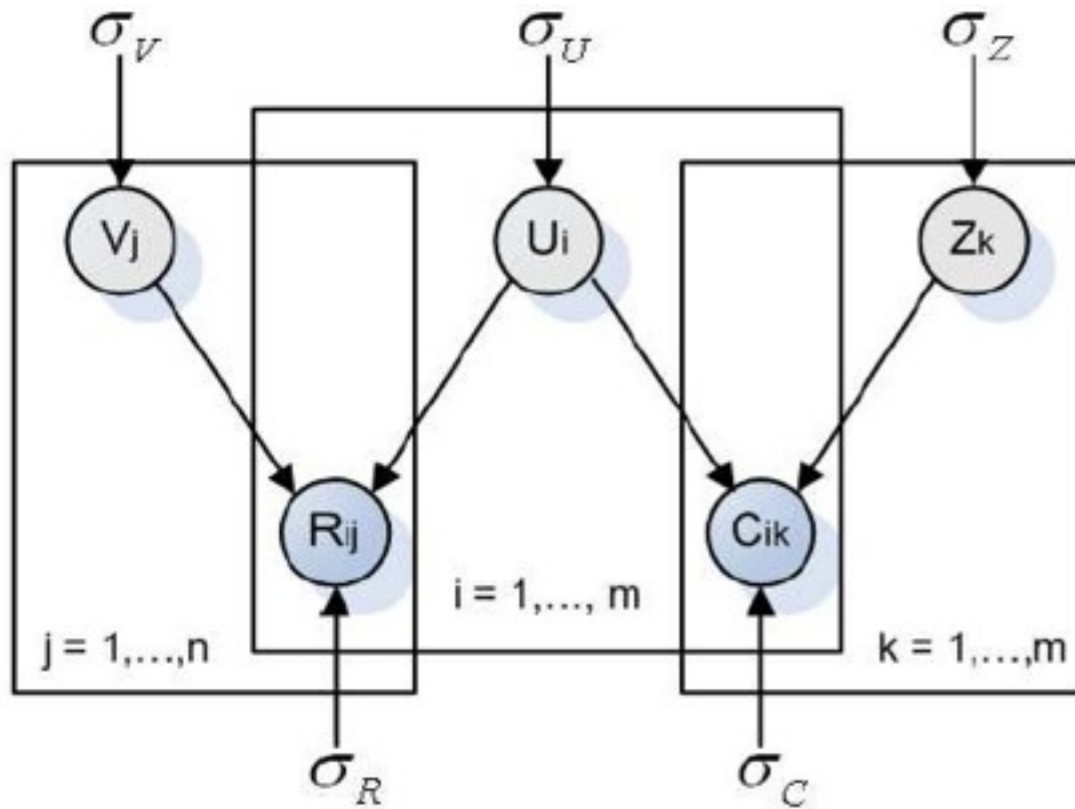
	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$
$u_1$		5	2		3	
$u_2$	4			3		4
$u_3$			2			2
$u_4$	5			3		
$u_5$		5	5			3



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

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

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

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





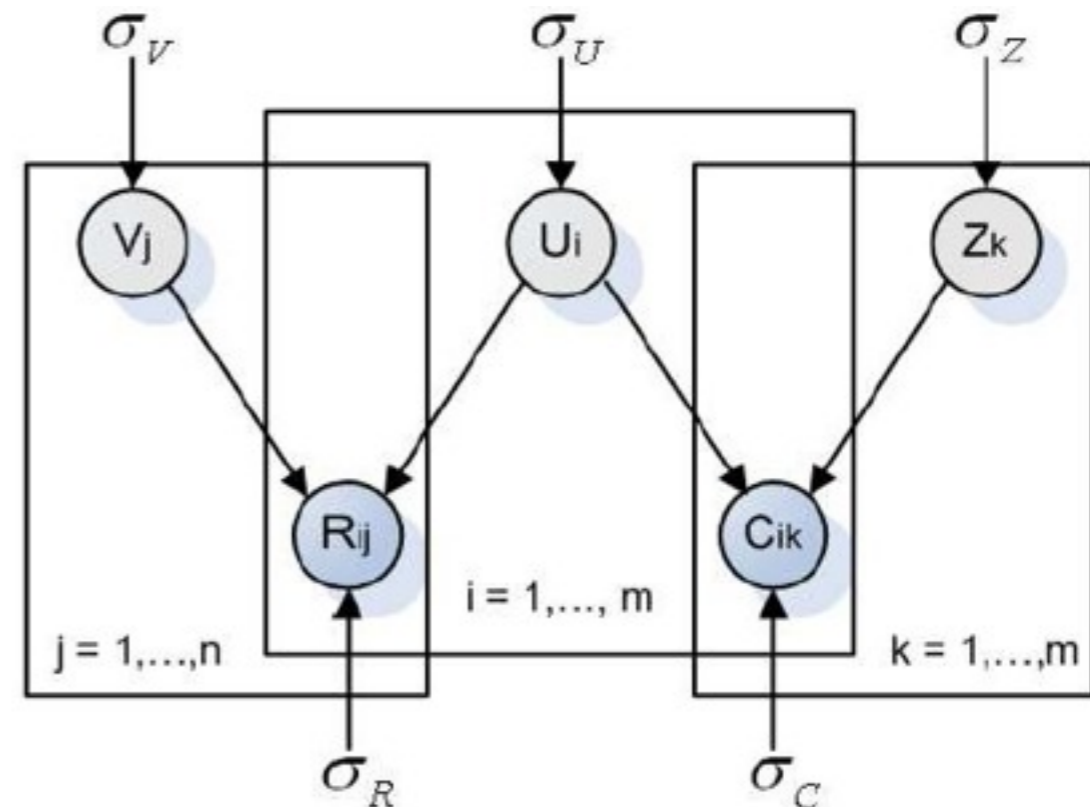
# Complexity Analysis

- 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]



# 1<sup>st</sup> Motivation





# 1<sup>st</sup> Motivation



Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA



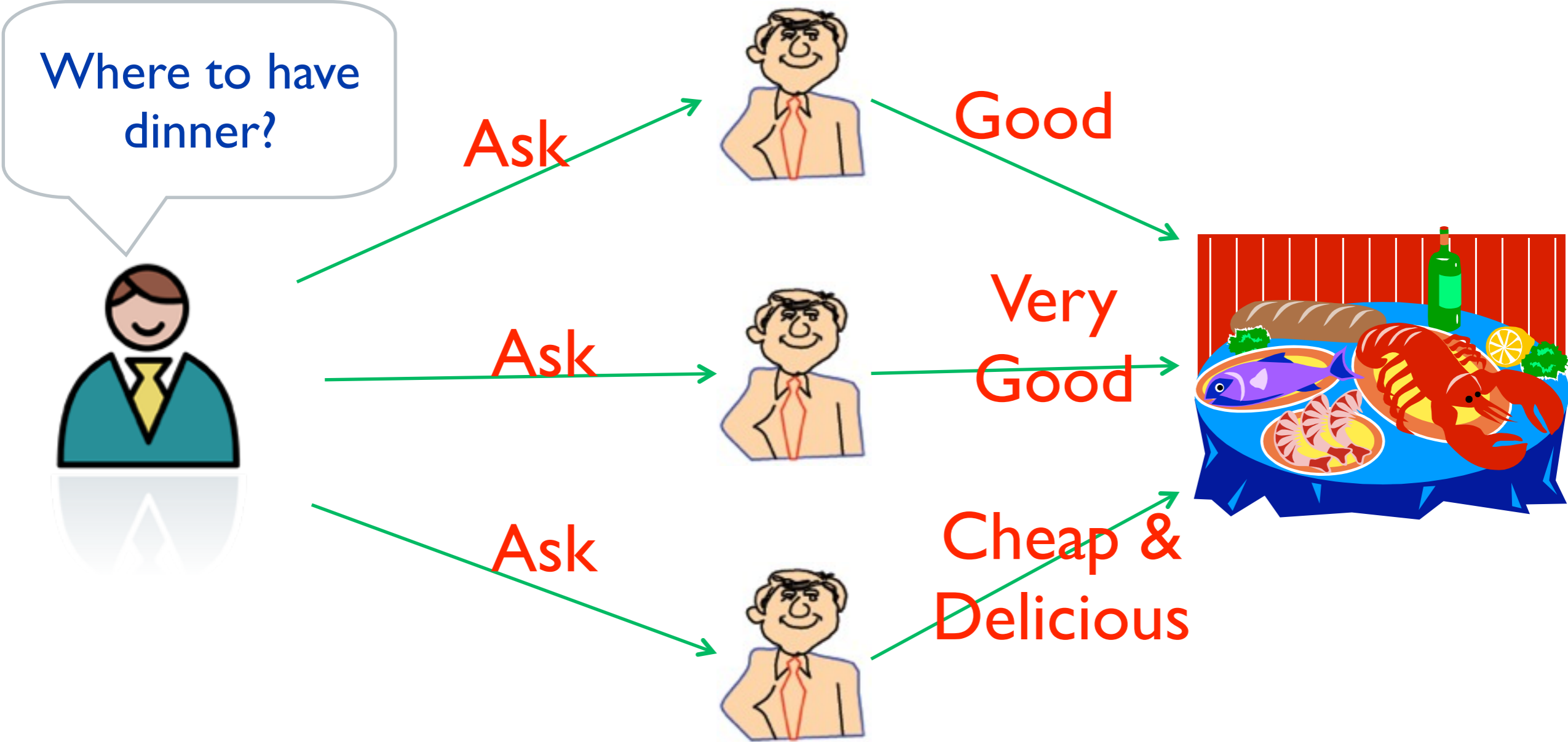


# 1<sup>st</sup> Motivation

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

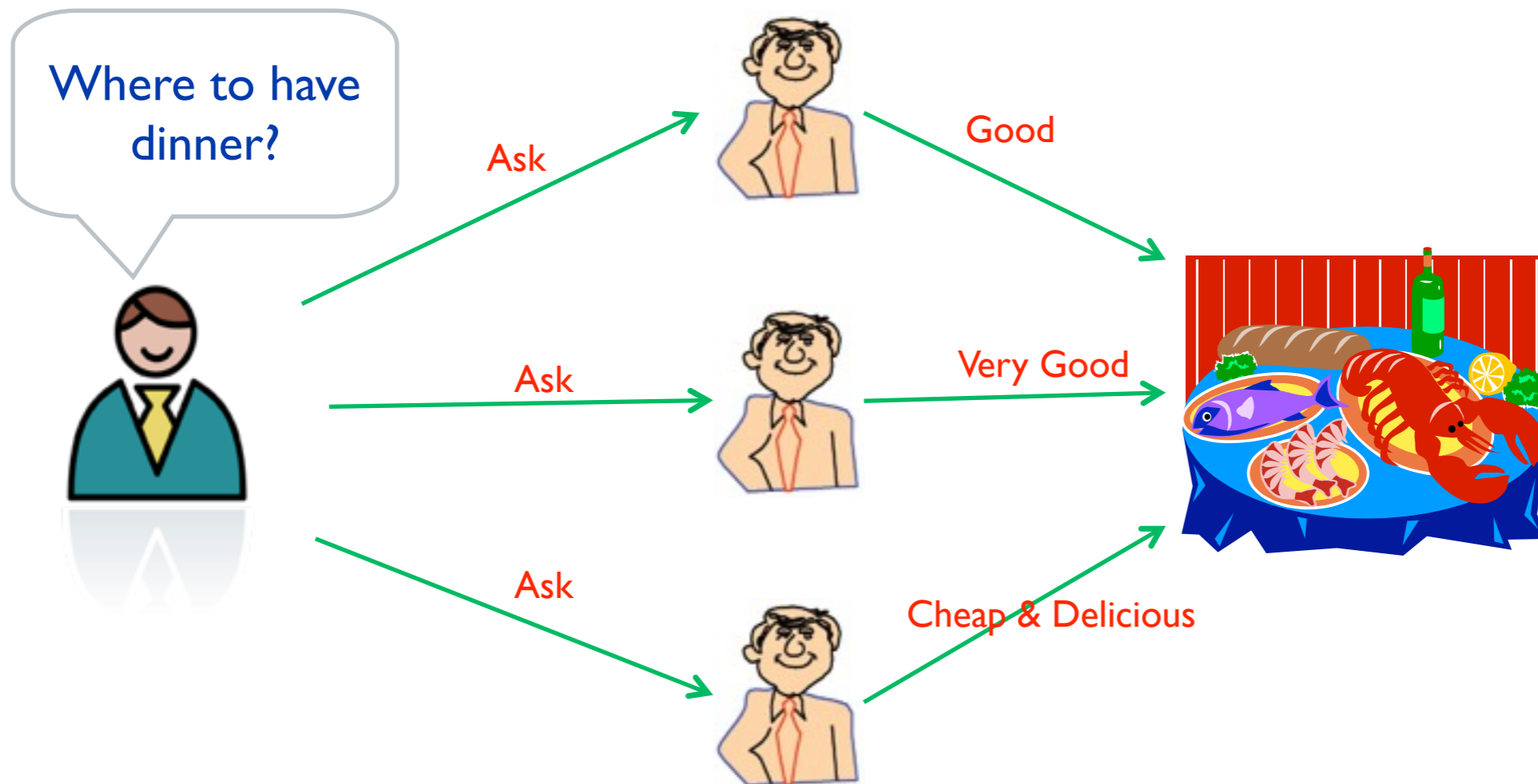


# 2<sup>nd</sup> Motivation



# 2<sup>nd</sup> Motivation

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





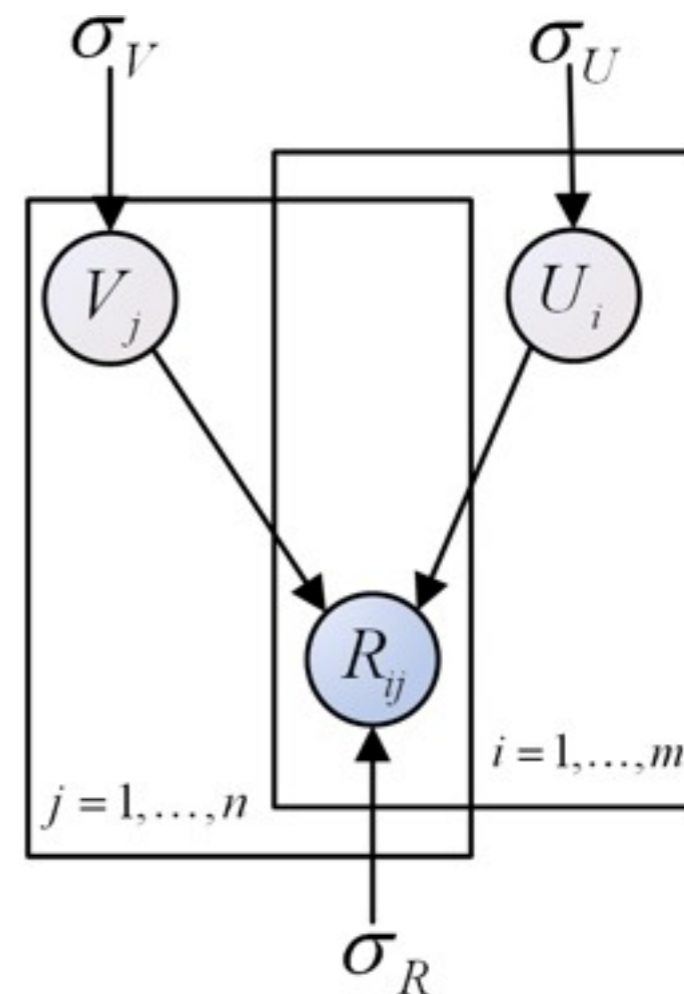
# Motivations

- 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]



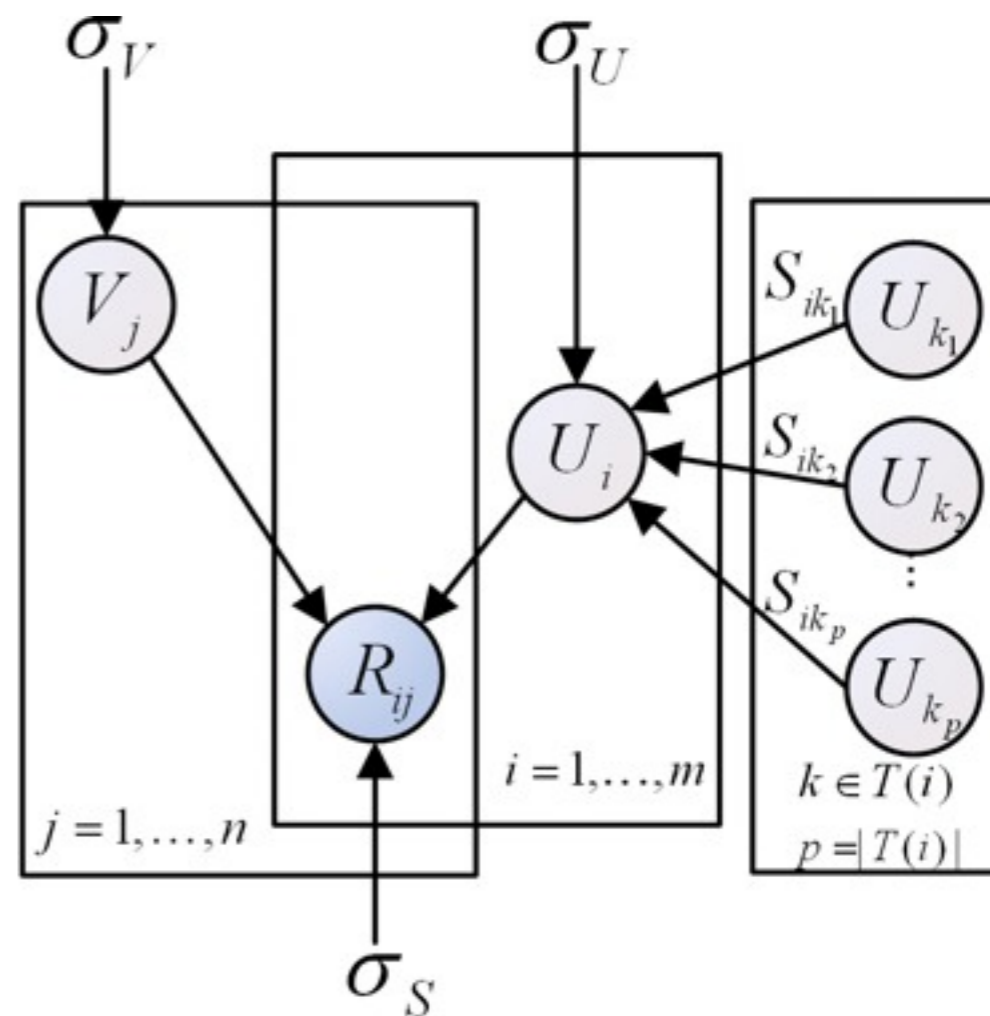
# Recommendations by Trusted Friends

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

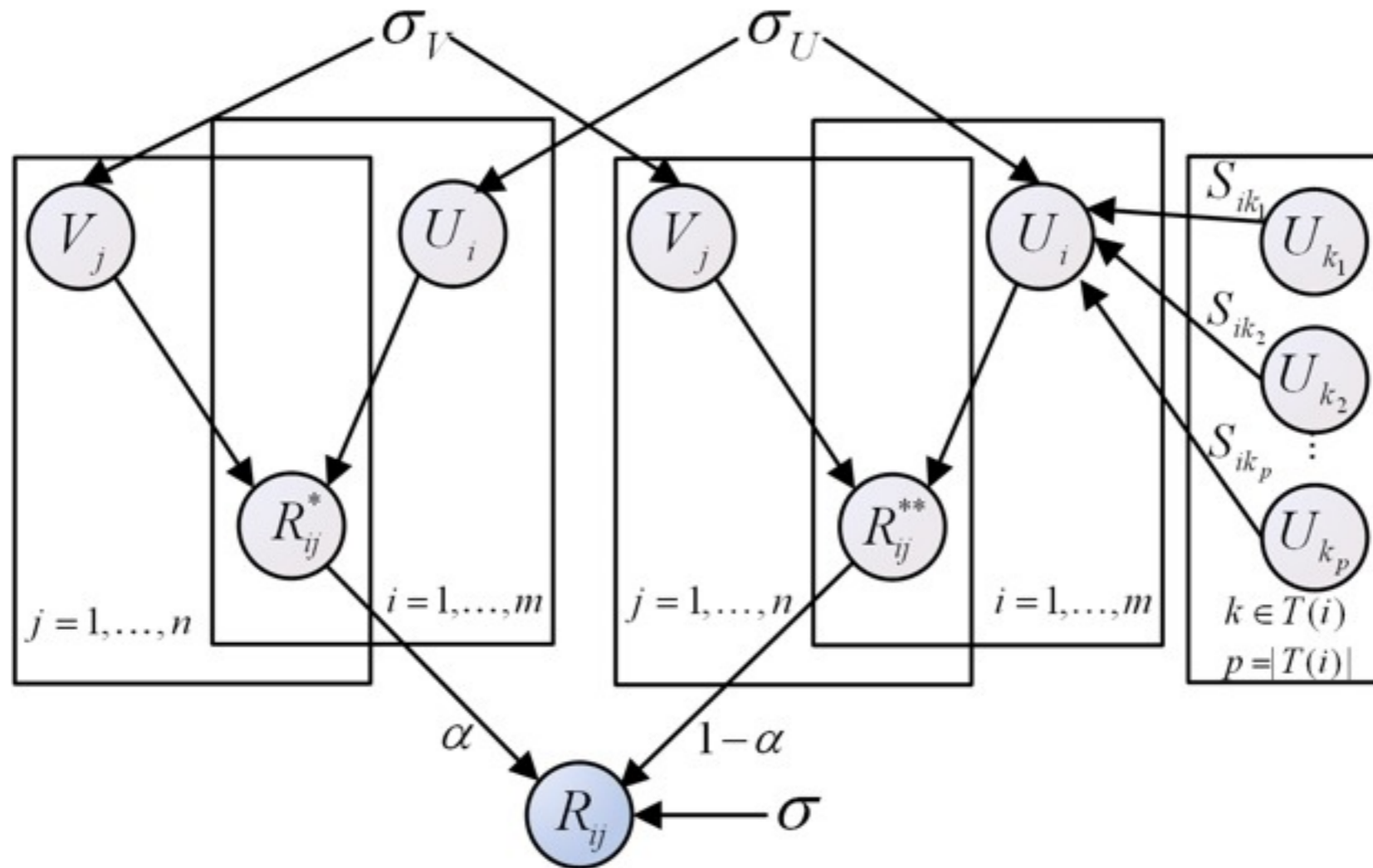
$$\hat{R}_{ik} = \sum_{j \in \mathcal{T}(i)} R_{jk} S_{ij}$$

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

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



# 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	<b>0.8377</b>
	RMSE	1.1688	1.2375	1.1649	1.1575	1.1959	1.1333	<b>1.1109</b>
80%	MAE	0.9285	0.9913	0.8975	0.8951	0.9221	0.8638	<b>0.8594</b>
	RMSE	1.1817	1.2584	1.1861	1.1826	1.2140	1.1530	<b>1.1346</b>

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	<b>0.8367</b>
	RMSE	1.1688	1.2375	1.1621	1.1544	1.1917	1.1293	<b>1.1094</b>
80%	MAE	0.9285	0.9913	0.8951	0.8886	0.9215	0.8580	<b>0.8537</b>
	RMSE	1.1817	1.2584	1.1832	1.1760	1.2132	1.1492	<b>1.1256</b>

**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)

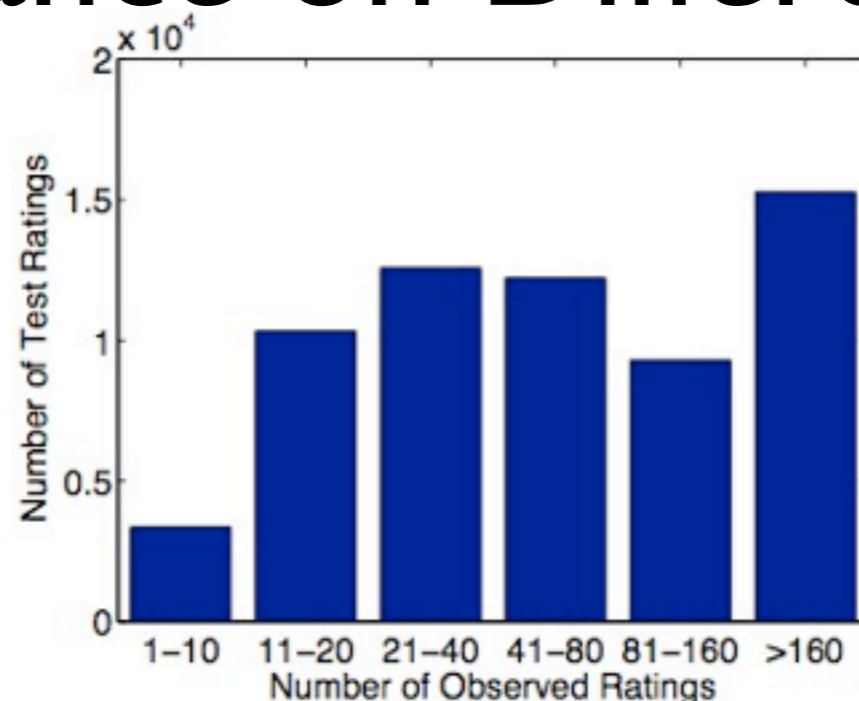


# Performance on Different Users

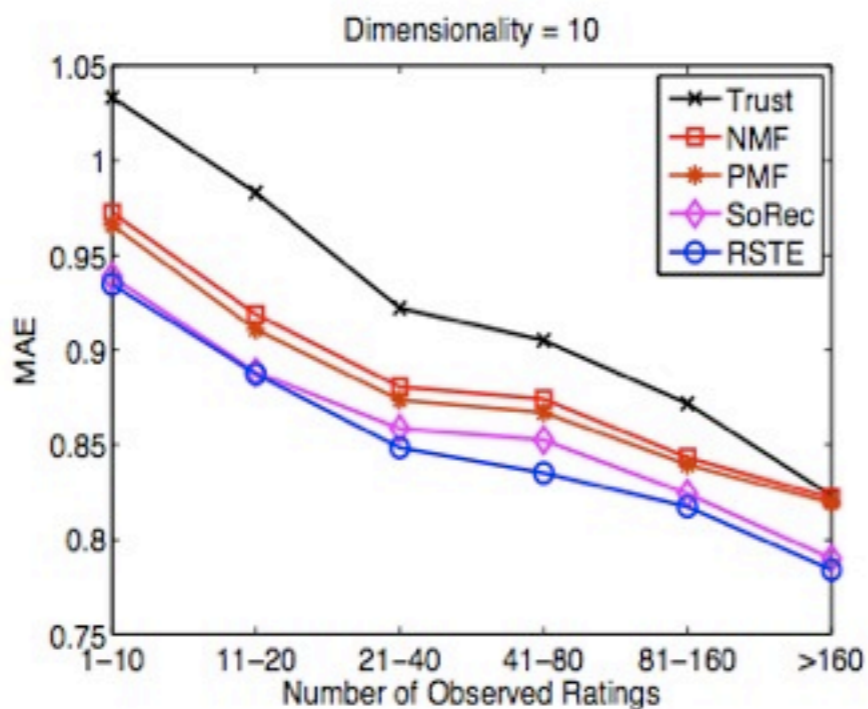
- Group all the users based on the number of observed ratings in the training data
- 6 classes: “1 – 10”, “11 – 20”, “21 – 40”, “41 – 80”, “81 – 160”, “> 160”,



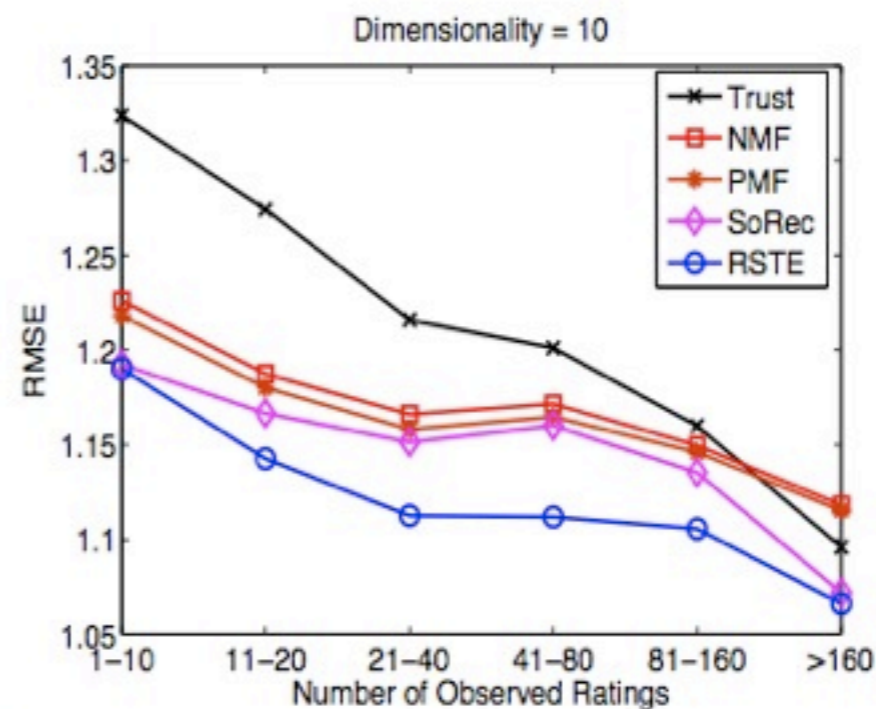
# Performance on Different Users



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



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

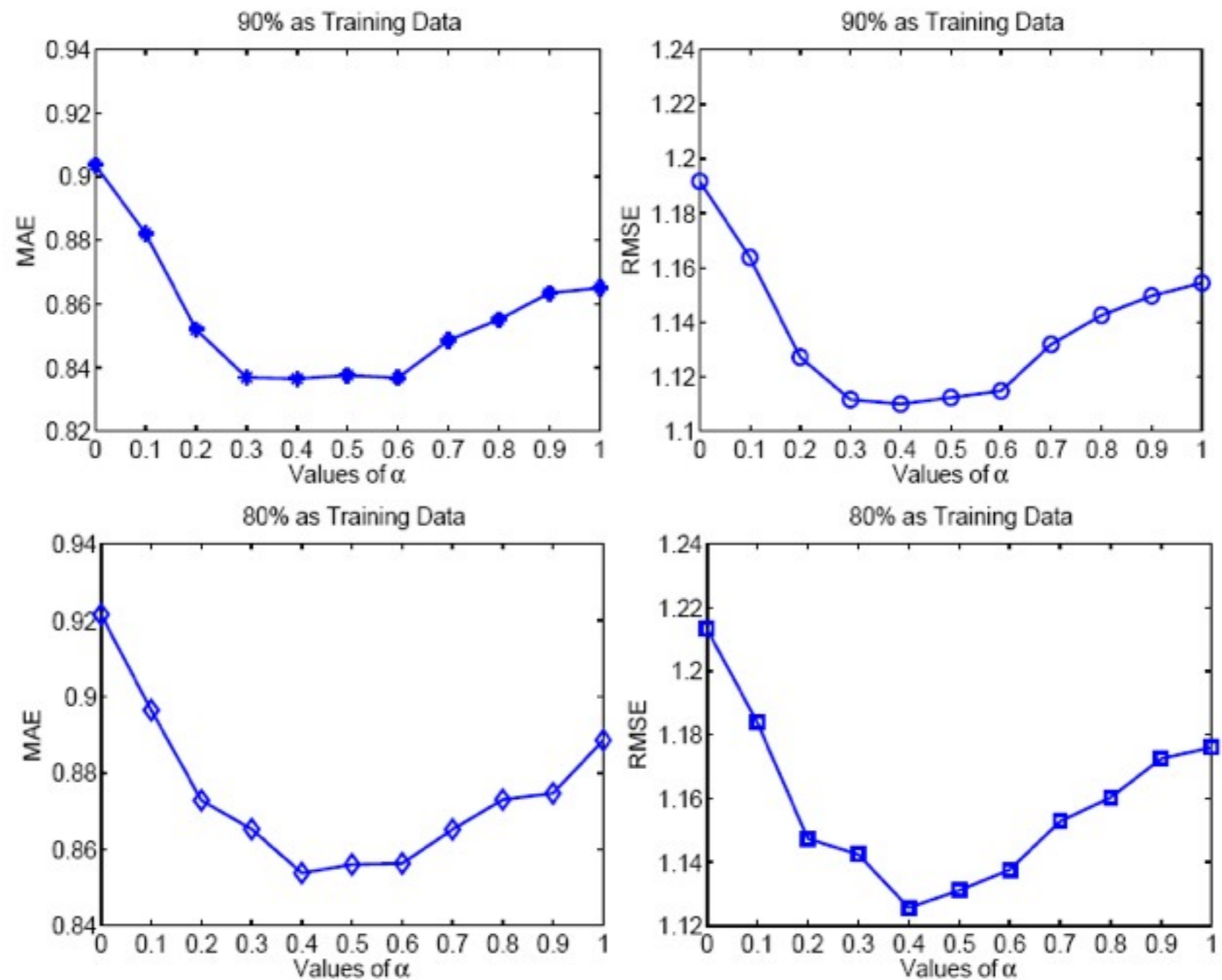


(c) RMSE Comparison on Different User Rating Scales (90% as Training Data)





# Impact of Parameter Alpha

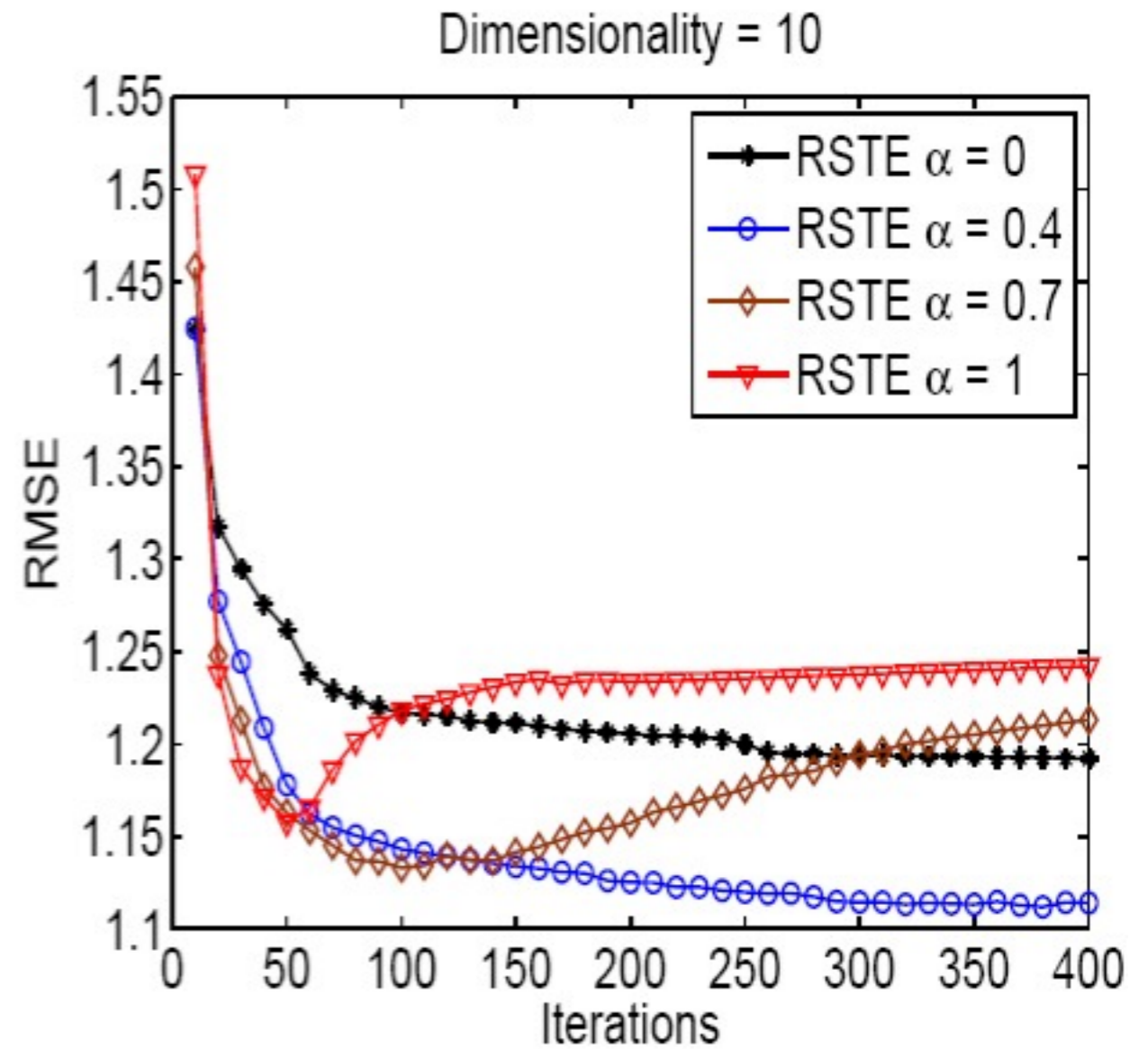
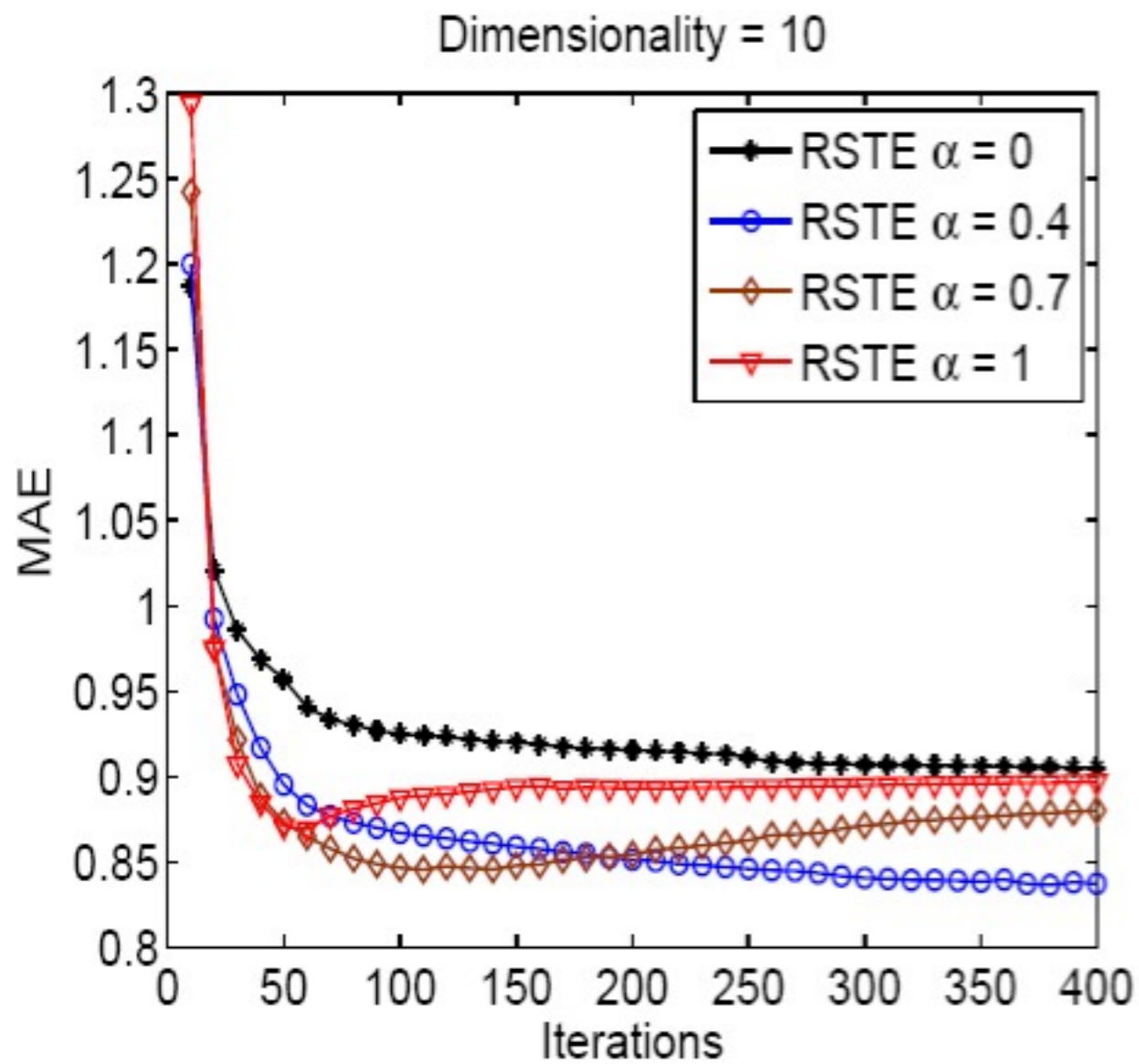


Impact of Parameter  $\alpha$  (Dimensionality = 10)





# MAE and RMSE Changes with Iterations

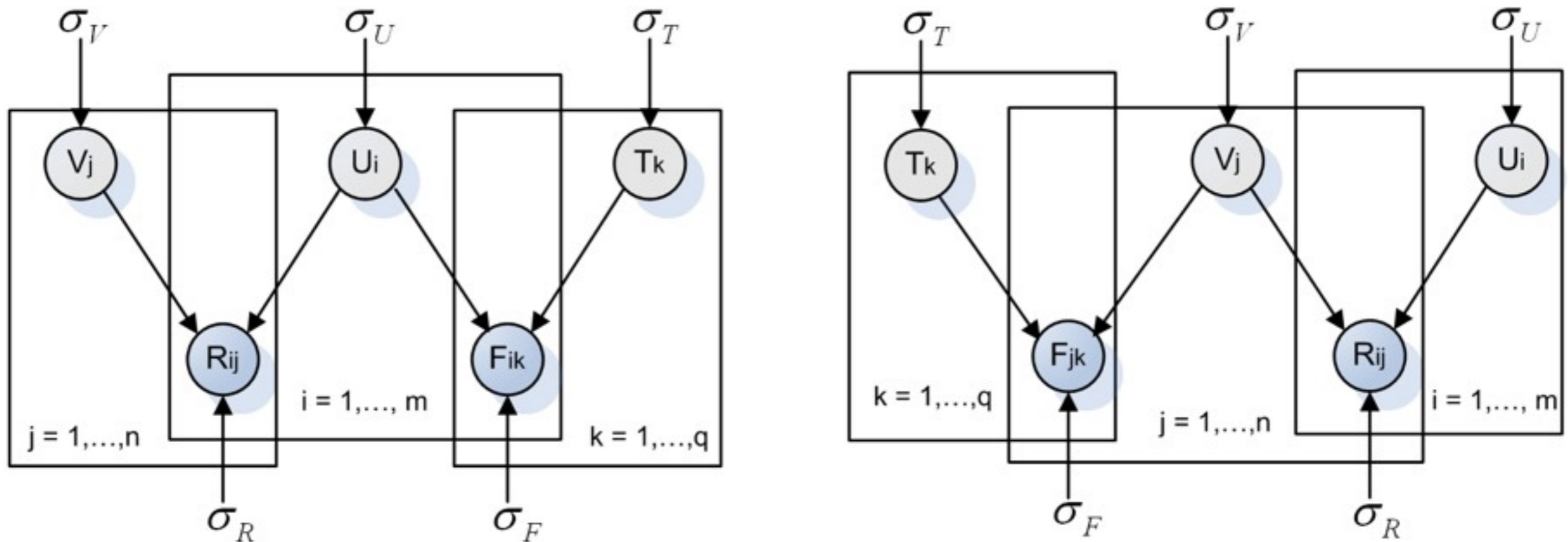


90% as Training Data



# 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	<b>0.6199</b>	<b>0.6407</b>	<b>0.6395</b>	<b>0.7026</b>
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	<b>0.6187</b>	<b>0.6395</b>	<b>0.6584</b>	<b>0.7016</b>





# Further Discussion of SoRec

- RMSE

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

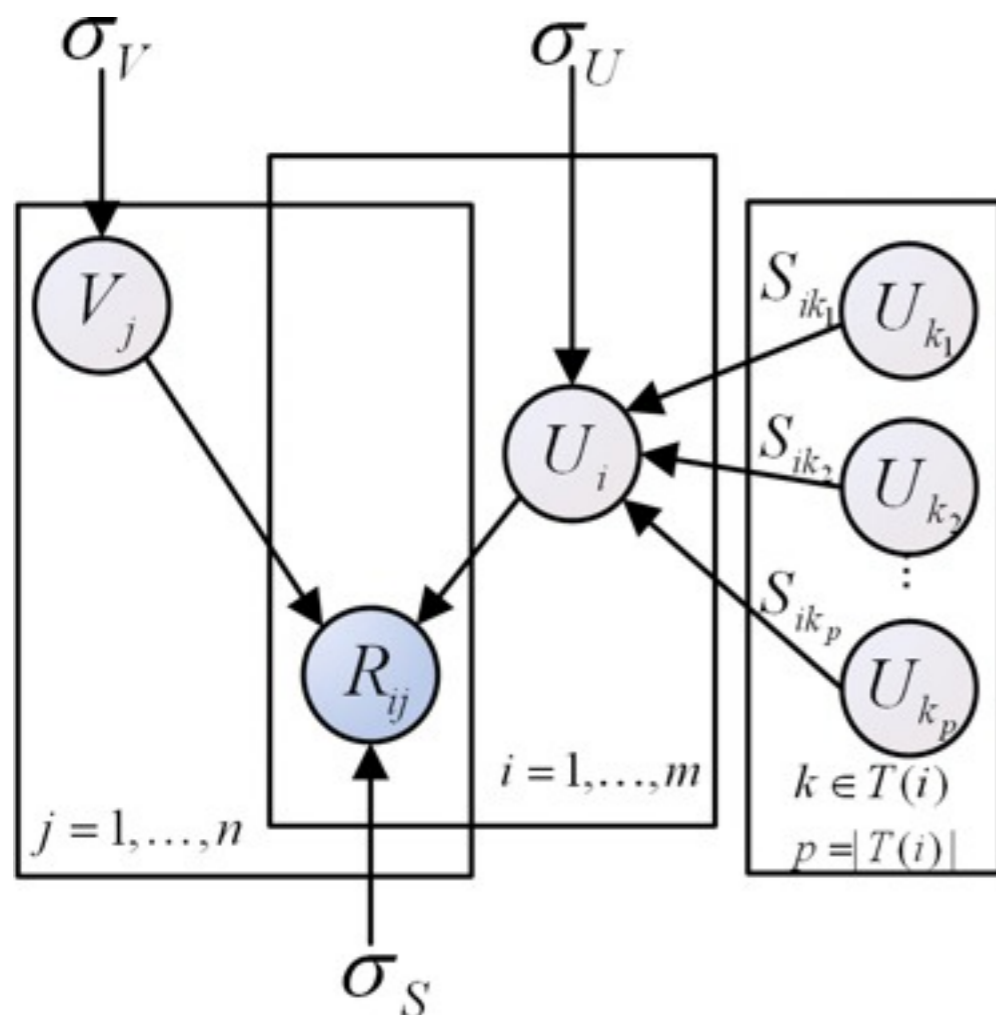
Methods		80% Training	50% Training	30% Training	10% Training
User Mean		0.9779	0.9816	0.9869	1.1587
Item Mean		0.9440	0.9463	0.9505	0.9851
5D	SVD	0.8327	0.8524	0.8743	0.9892
	PMF	0.8310	0.8582	0.8758	0.9698
	SoRecUser	0.8121	0.8384	0.8604	0.9042
	SoRecItem	<b>0.8112</b>	<b>0.8370</b>	<b>0.8591</b>	<b>0.9033</b>
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	<b>0.8097</b>	<b>0.8359</b>	<b>0.8578</b>	<b>0.9019</b>





# Further Discussion of RSTE

- Relationship with Neighborhood-based methods

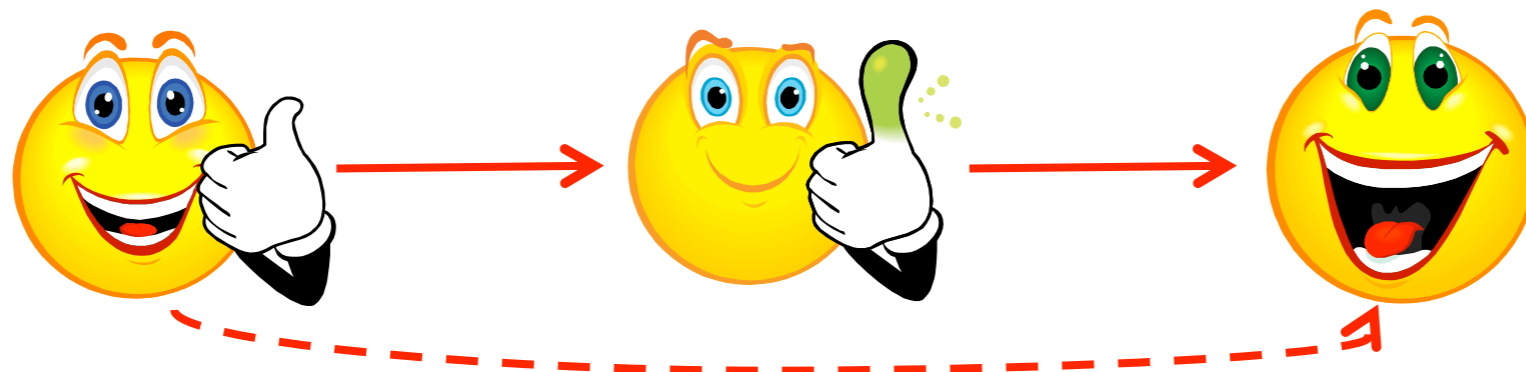


- The trusted friends are actually the explicit neighbors
- We can easily apply this method to include implicit neighbors
- Using PCC to calculate similar users for every user

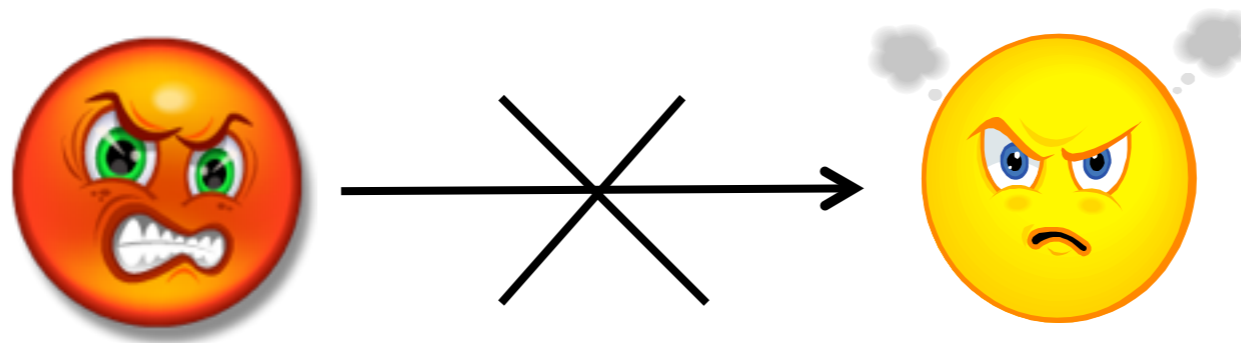


# What We Cannot Model Using SoRec and RSTE?

- 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  $U_t$ .



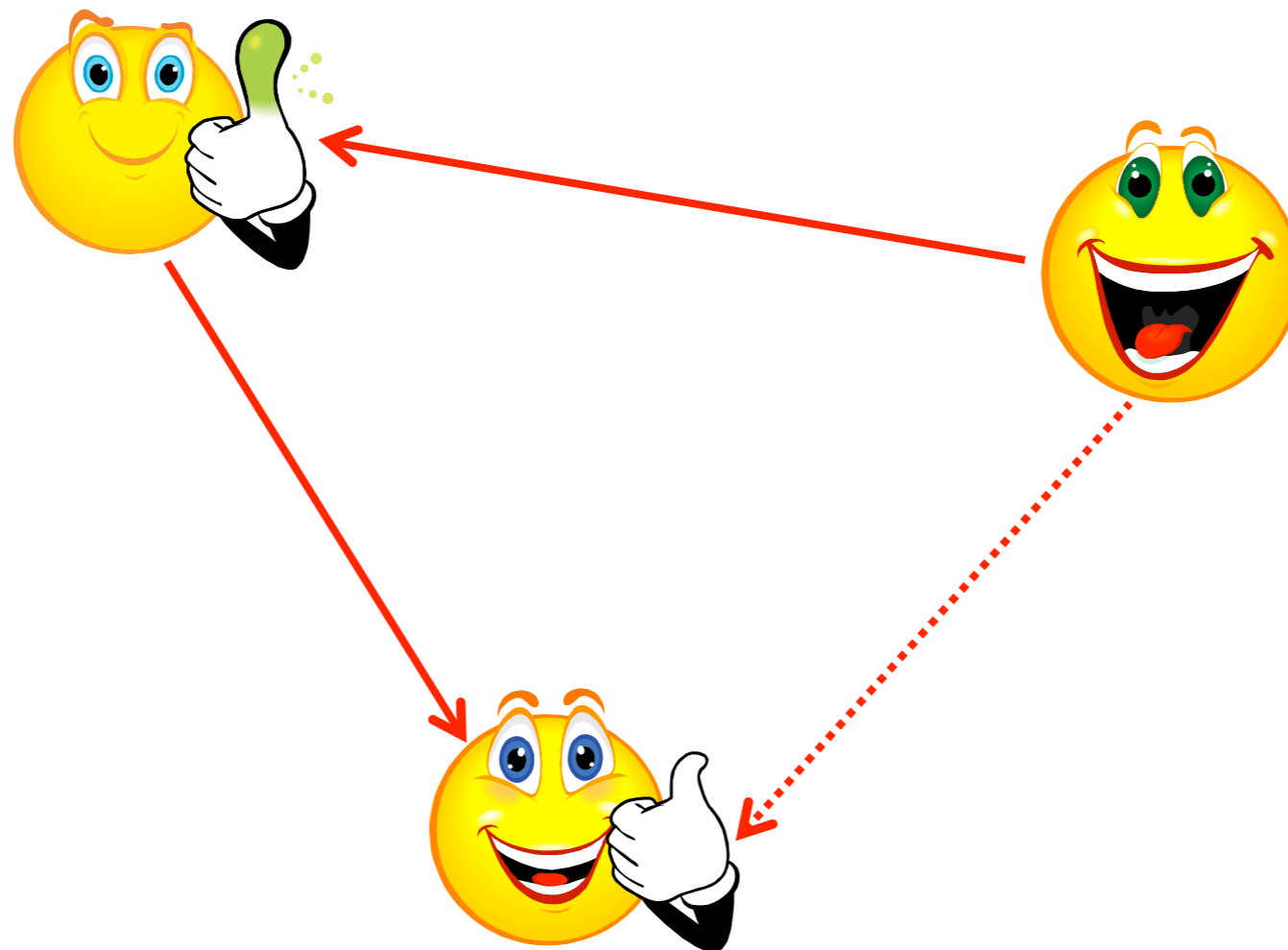
# Trust

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

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

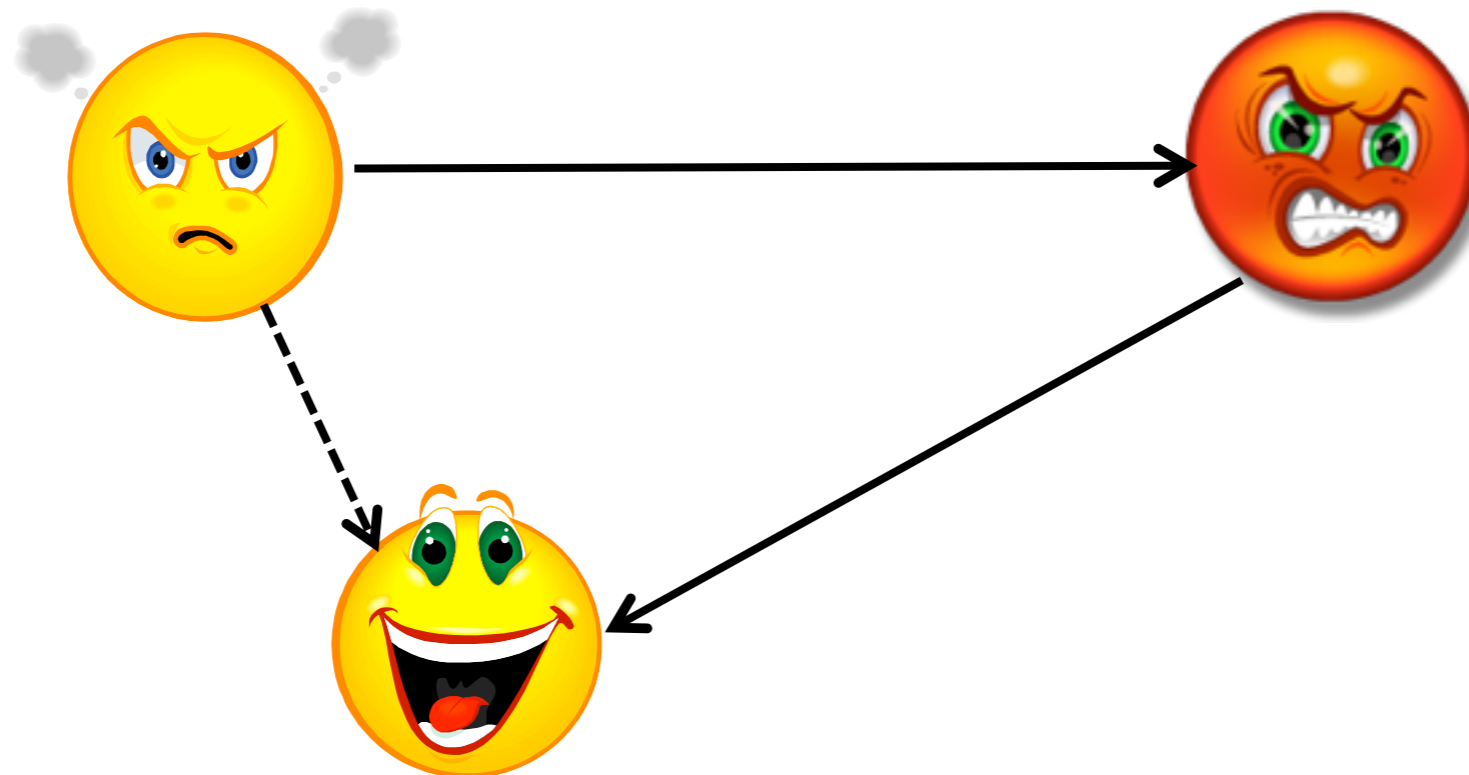
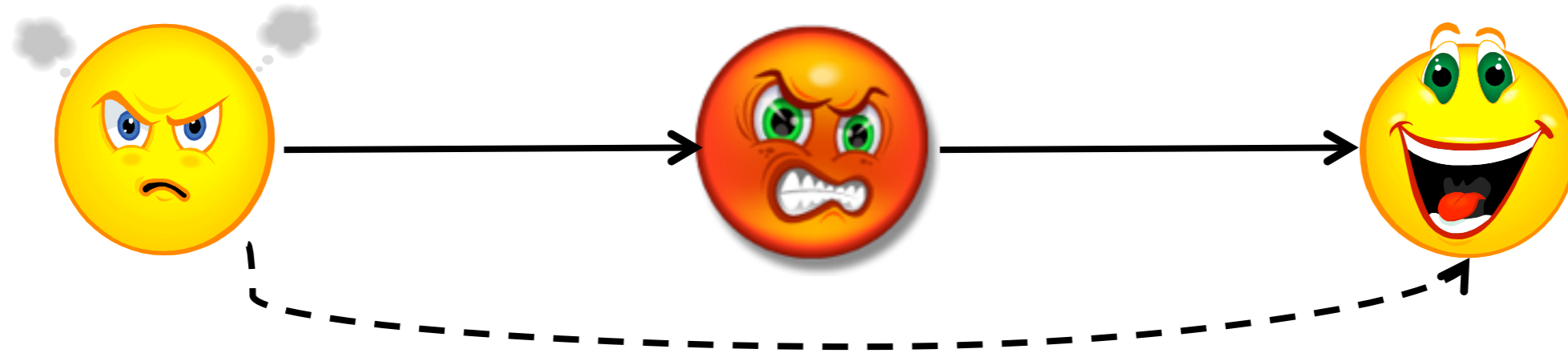


# Trust Propagation





# 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	<b>1.177</b>
		10D	1.214	1.198	1.185	<b>1.176</b>
	10%	5D	0.990	0.944	0.932	<b>0.924</b>
		10D	0.977	0.941	0.931	<b>0.923</b>
	20%	5D	0.819	0.788	0.723	<b>0.721</b>
		10D	0.818	0.787	0.723	<b>0.720</b>





# Impact of Parameters

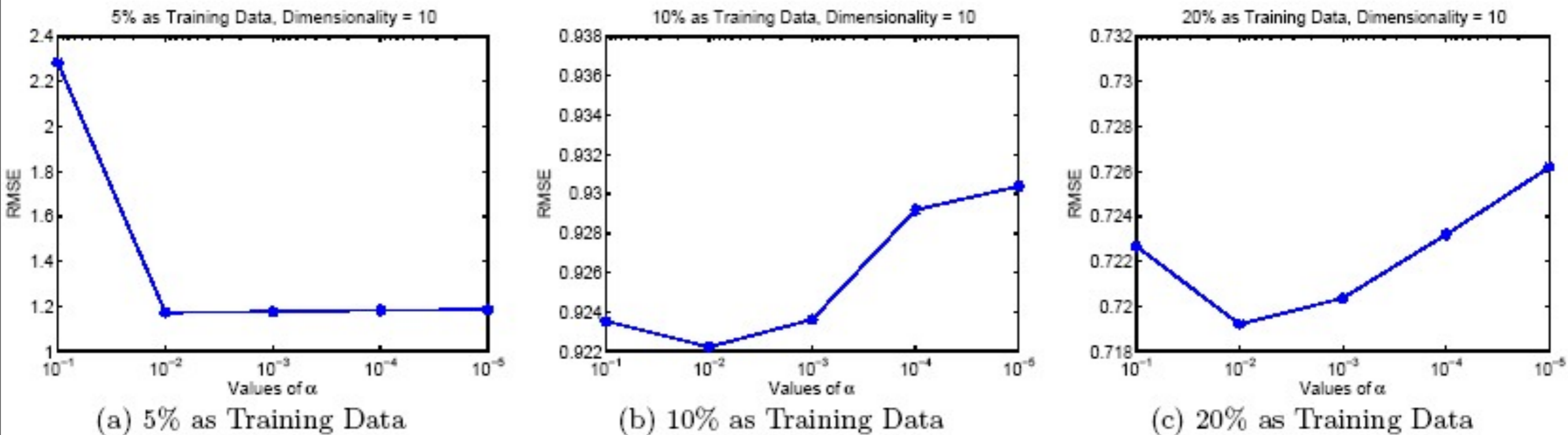


Figure 6: Impact of Parameter  $\alpha$

**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

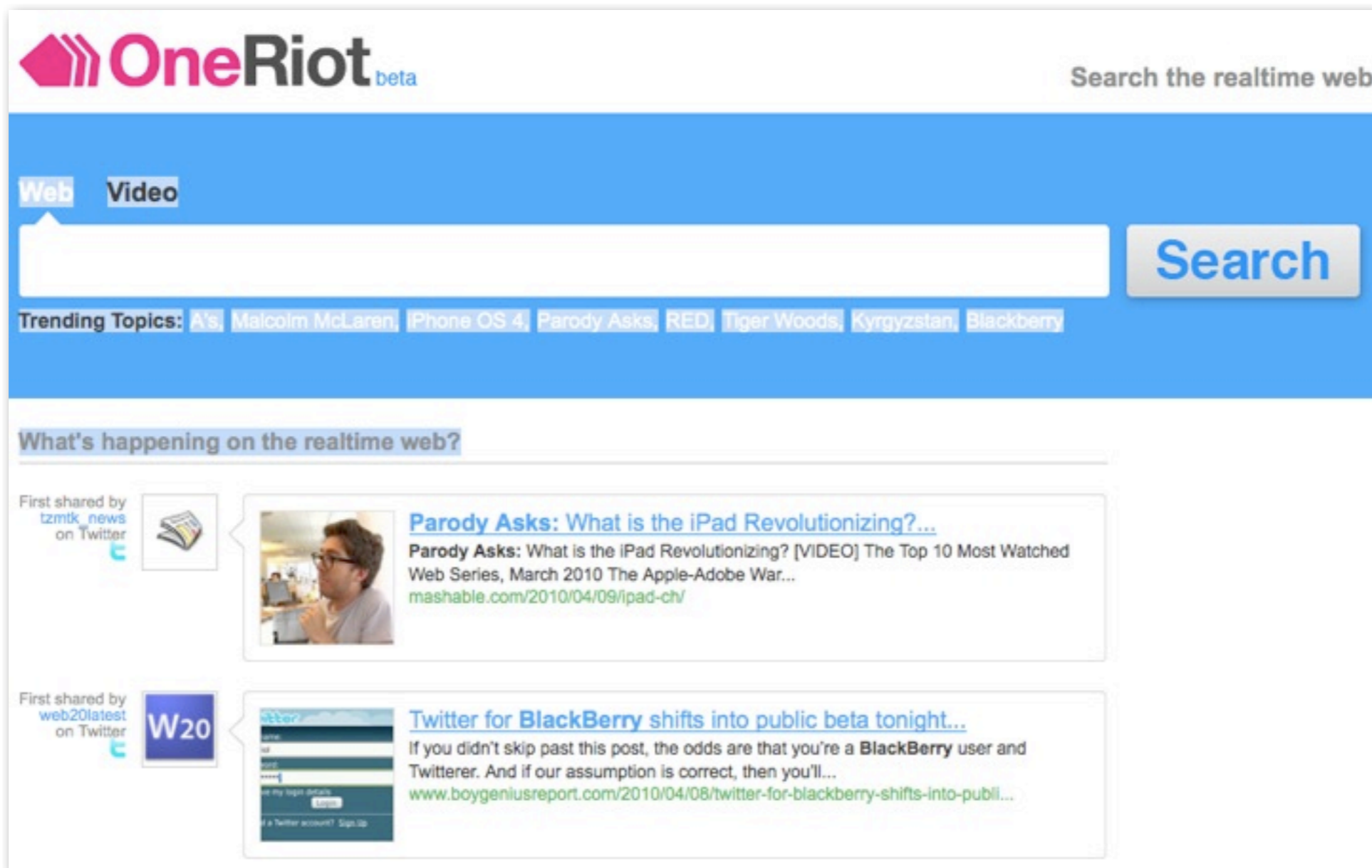
- Pulse Rank - OneRiot
- Reddit Algorithm
- Digg Algorithm
- Google's Page Rank





# 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 displays 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

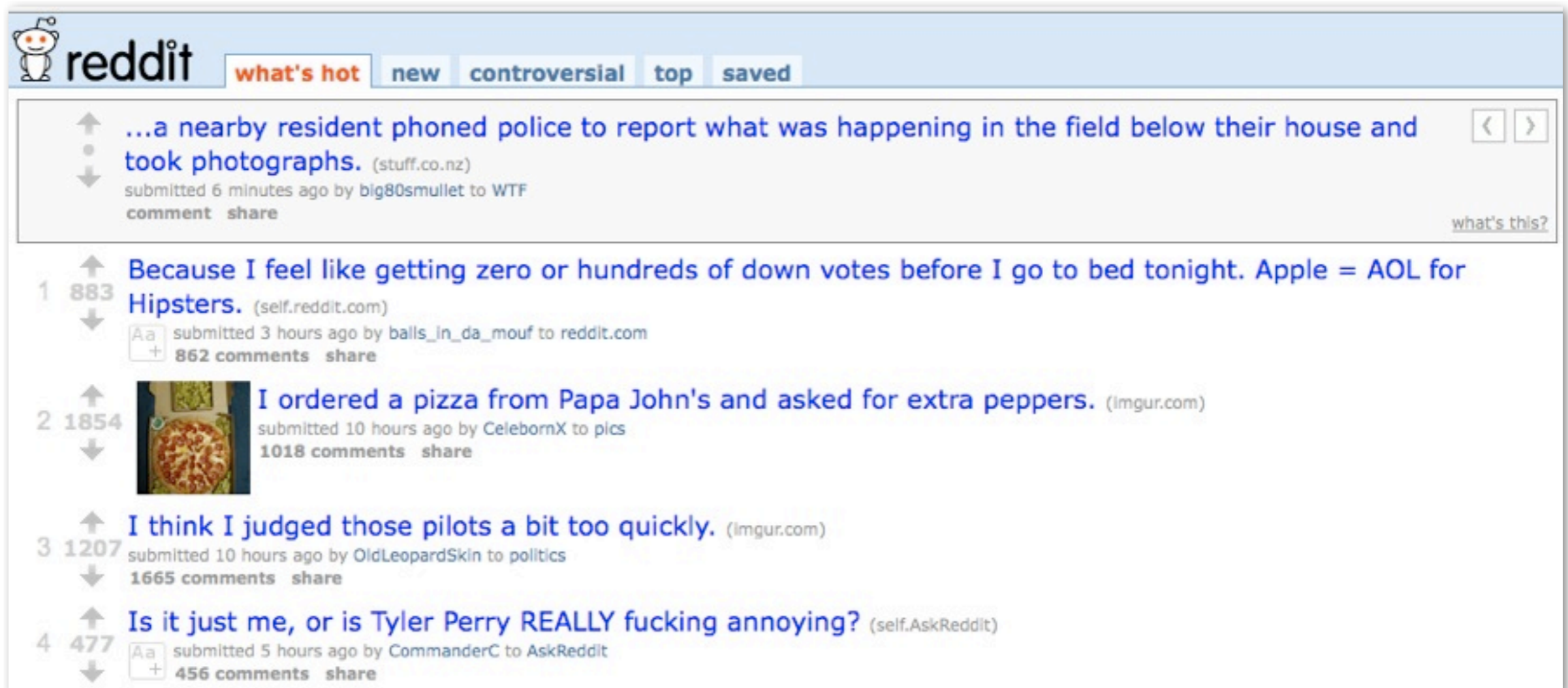
- “Pulse Rank” algorithm looks at dozens of factors that give “weight” to certain results
- **Freshness:** Is the most recently published content necessarily the most relevant?
- **Domain Authority:** An article about Obama on New York Times should weight higher than the article on my blog.
- **People 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 interface. At the top, the Reddit logo is on the left, and navigation tabs for 'what's hot', 'new', 'controversial', 'top', and 'saved' are in the center. Below the navigation, a list of posts is displayed. Each post includes a rank number, an up/down arrow, a vote count, the post title, the source, the submitter's name, the time since submission, and the number of comments. The first post is partially cut off. The second post is titled 'Because I feel like getting zero or hundreds of down votes before I go to bed tonight. Apple = AOL for Hipsters.' with 883 votes. The third post is titled 'I ordered a pizza from Papa John's and asked for extra peppers.' with 1854 votes and includes a small image of a pizza. The fourth post is titled 'I think I judged those pilots a bit too quickly.' with 1207 votes. The fifth post is titled 'Is it just me, or is Tyler Perry REALLY fucking annoying?' with 477 votes.



# 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

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

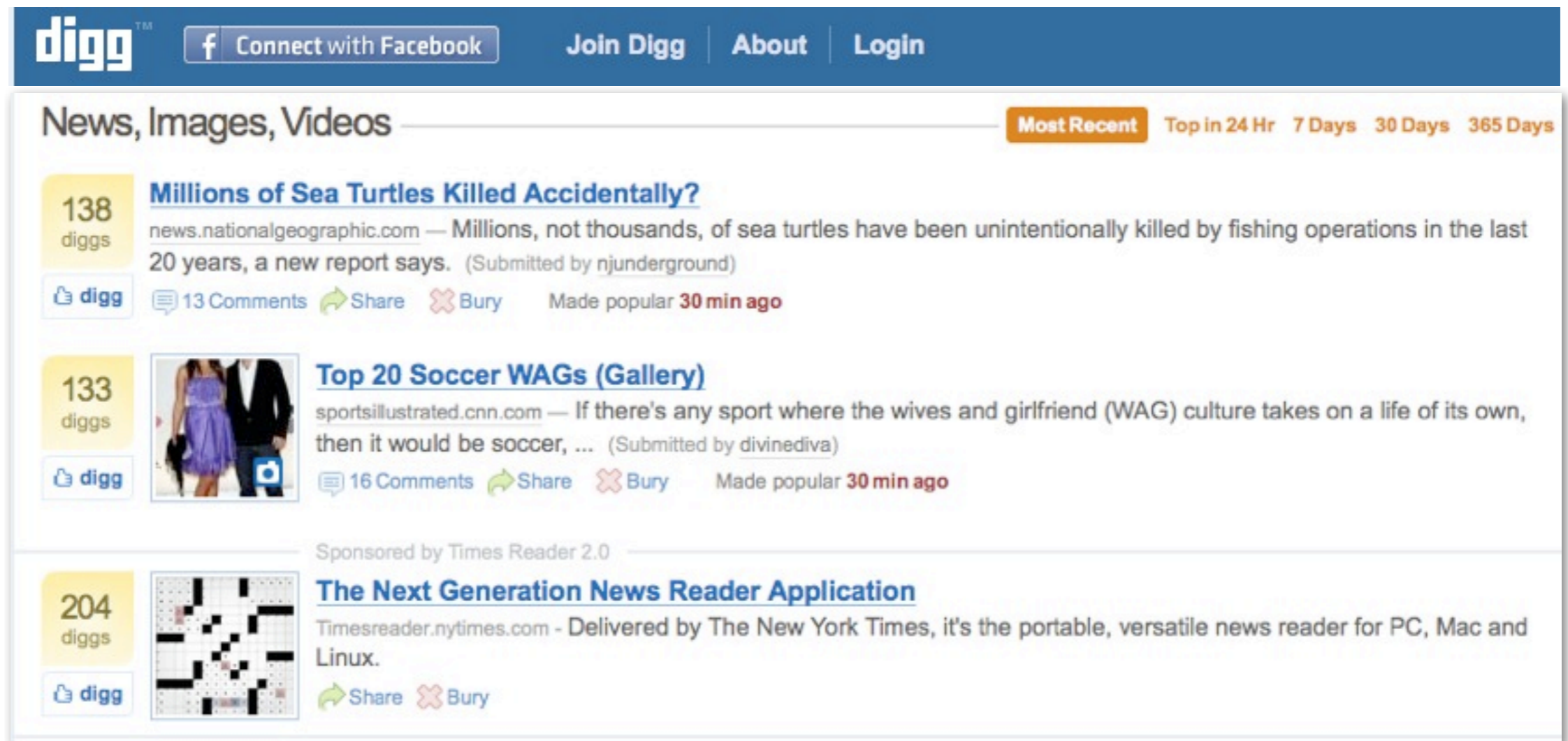
From <http://uggedal.com/reddit.cf.algorithm.png>







# 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?](#)  
news.nationalgeographic.com — 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)  
13 Comments | Share | Bury | Made popular 30 min ago
- 133 diggs** |  | [Top 20 Soccer WAGs \(Gallery\)](#)  
sportsillustrated.cnn.com — 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)  
16 Comments | Share | Bury | Made popular 30 min ago
- 204 diggs** |  | [The Next Generation News Reader Application](#)  
Timesreader.nytimes.com - Delivered by The New York Times, it's the portable, versatile news reader for PC, Mac and Linux.  
Share | Bury

A sponsored message 'Sponsored by Times Reader 2.0' is visible between the second and third news items.



# 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.

From <http://www.seopedia.org/tips-tricks/social-media/the-digg-algorithm-unofficial-faq/>  
Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA



# 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

From <http://www.technologyreview.com/web/24353/?a=f>

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA





# 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



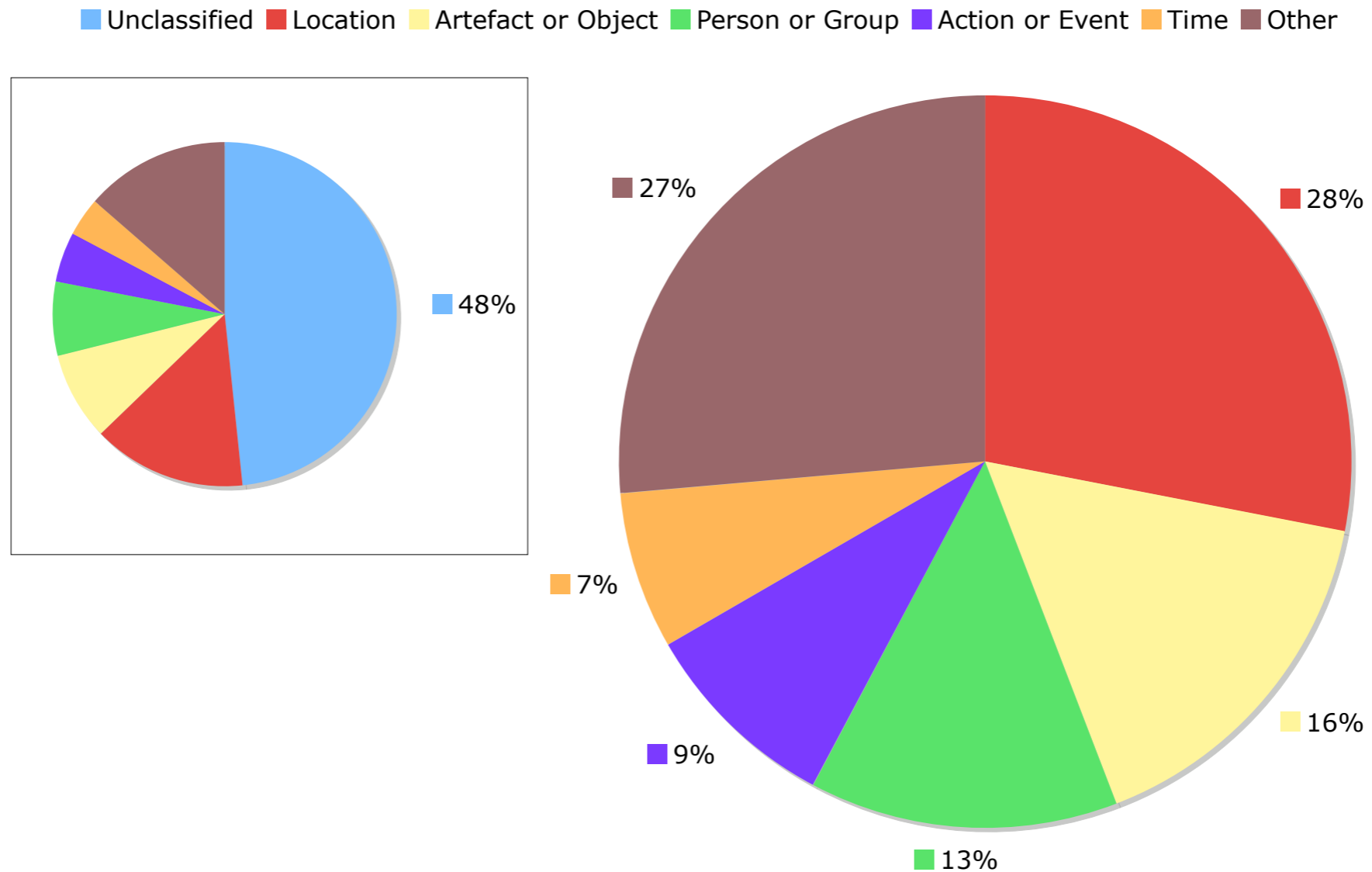
# Why need Tag Recommendation?

- 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]

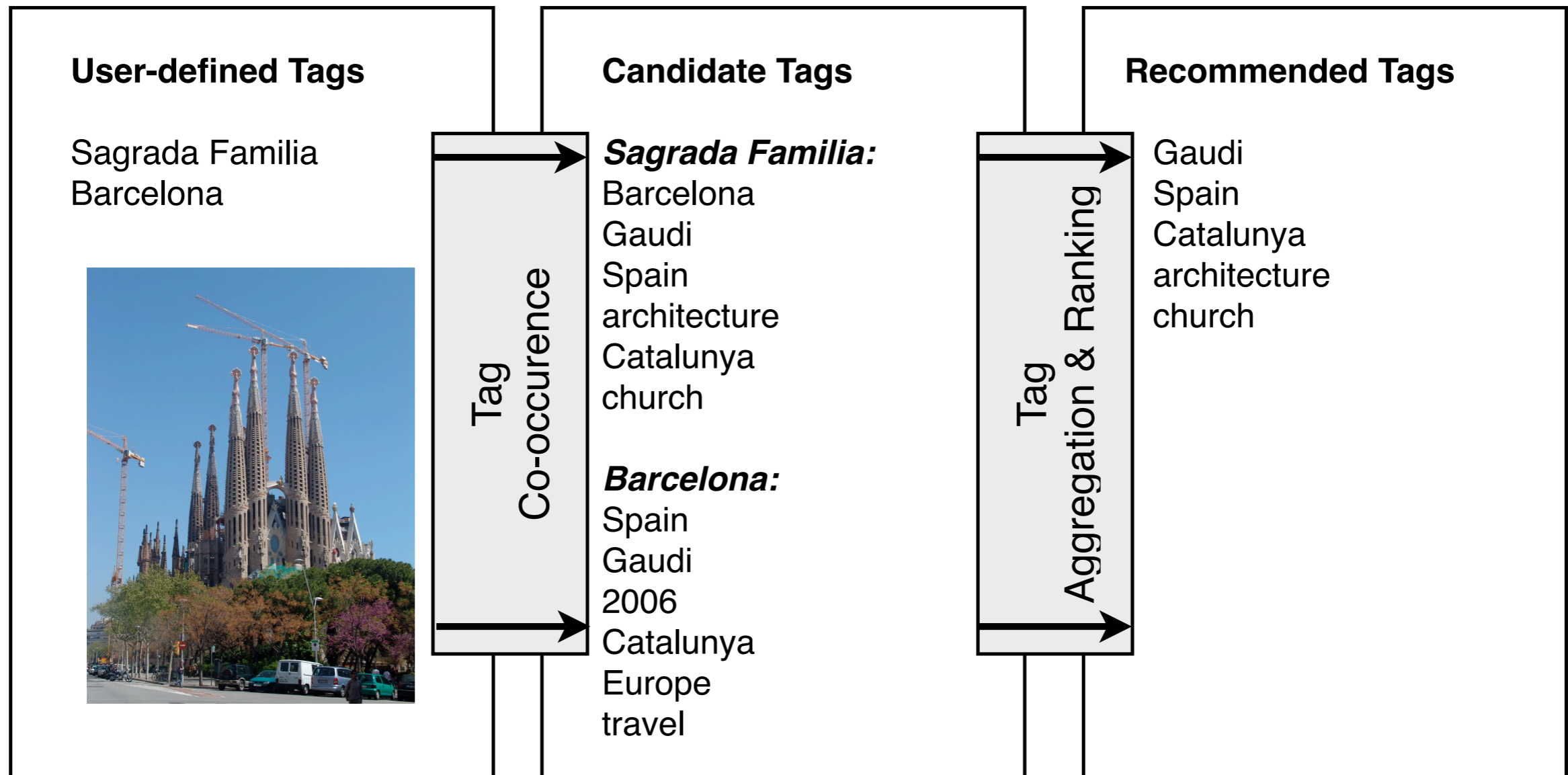
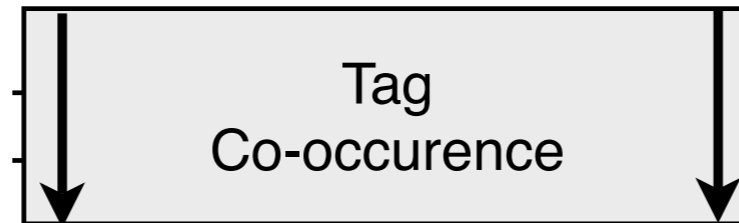


Figure 4: System overview of the tag recommendation process.



# 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**

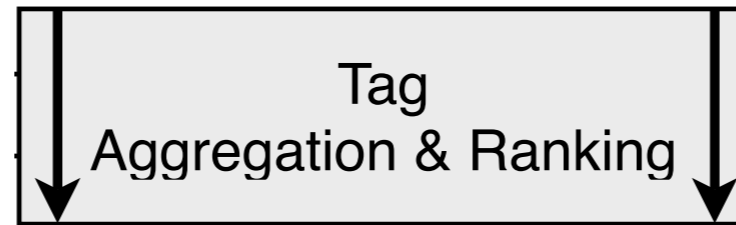
Good at suggesting  
diverse tags



# 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](#)  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](#)  Al

[Political Standoff in Bangkok Intensifies](#) New Yo

[Times Online](#) - [Reuters](#) - [The Associated Press](#)

[Wikipedia: National United Front of Democracy A](#)

[all 2,174 news articles](#) » [Email this story](#)



## 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

Google news



# 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]

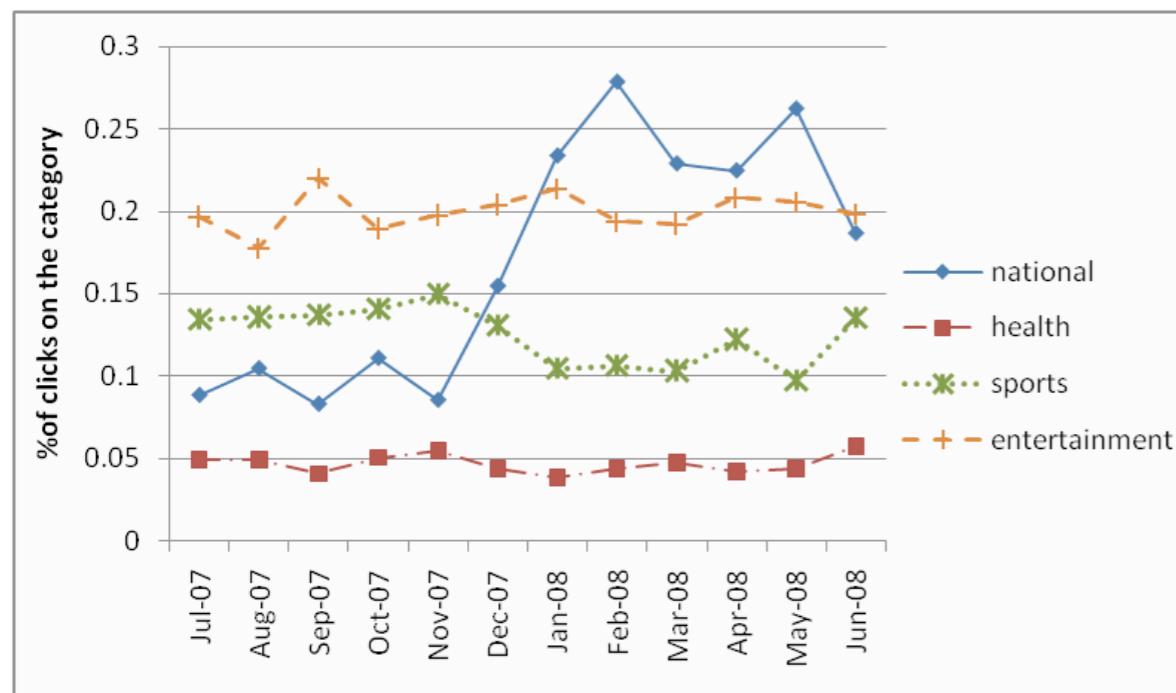


Figure 2. Interest distribution of US users over time

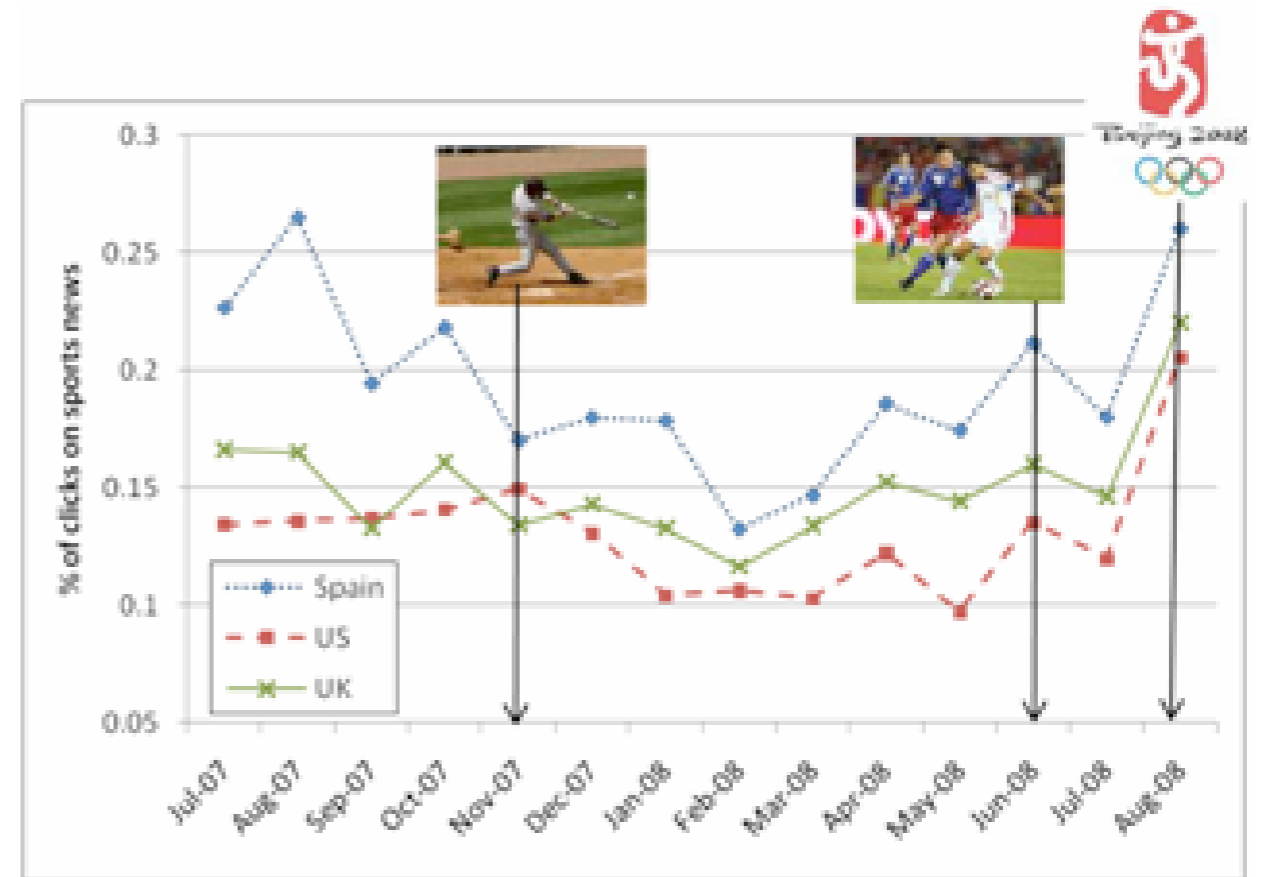


Figure 3. Change of interests in sports news over time



# 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:

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

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

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



# 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]

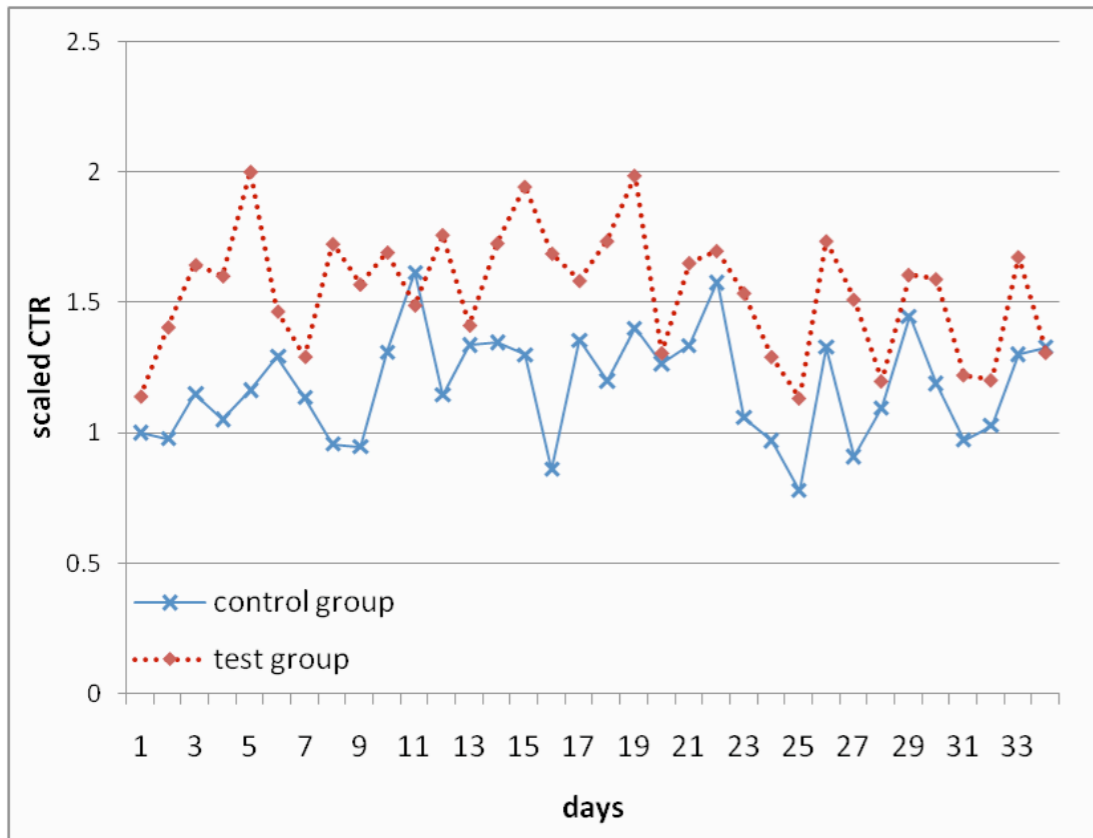


Figure 4. CTR of the recommended news section

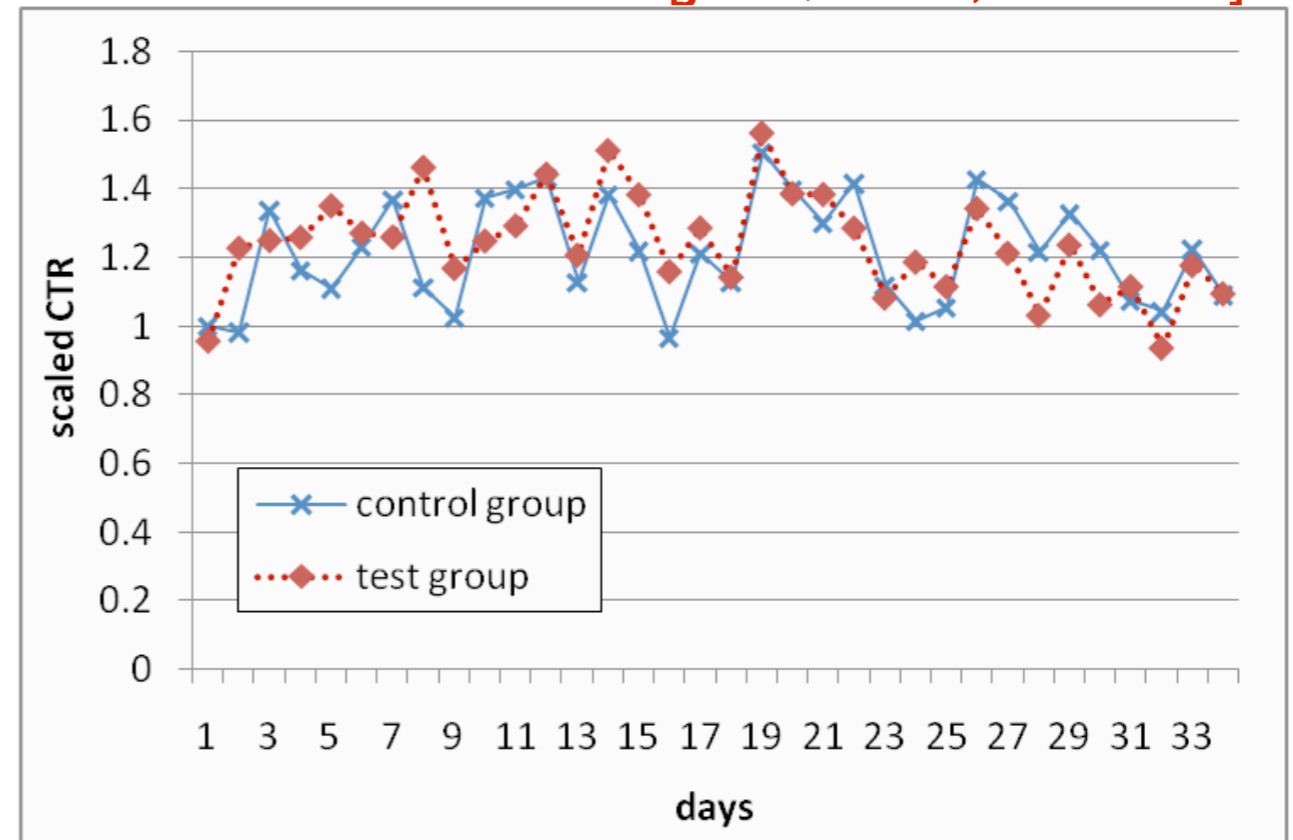


Figure 5. CTR of the Google News homepage

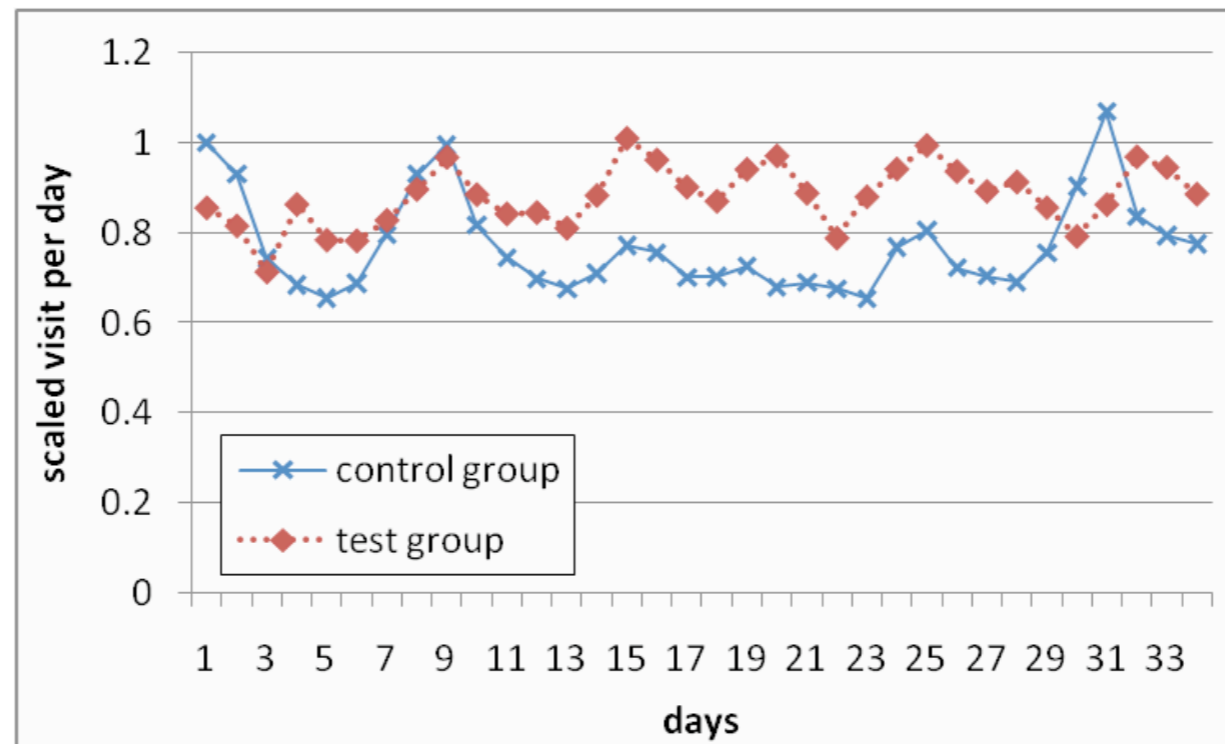


Figure 6. Frequency of website visit per day





# 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 'Suggestions' interface. At the top, the Facebook logo and a search bar are visible. Below the header, the 'Suggestions' section is titled 'Suggestions' and includes the instruction: '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 an action button. The first four rows contain individual users, while the fifth row contains a public profile (BFU).

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



# User Recommendation

The screenshot shows the Twitter interface with the 'Look who else is here. Start following them!' section. The 'Entertainment' category is selected in the left sidebar. The 'Sources in Entertainment' section lists five recommended users, each with a profile picture, name, handle, location, bio, and a 'follow' button.

**Twitter** Home Profile Find People Settings Help Sign out

## Look who else is here. Start following them!

Browse Suggestions Find Friends Invite By Email Find On Twitter

Look who else is here! Follow the ones you like.

- Art & Design
- Books
- Business
- Charity
- Entertainment**
- Family
- Fashion
- Food & Drink
- Funny
- Health
- Music
- News
- Politics
- Science
- Sports
- Staff Picks
- Staff Picks for Haiti
- Technology
- Travel
- Twitter

### Sources in Entertainment

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



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

[Jilin Chen, et al., CHI2009]

- On social networking site, 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]

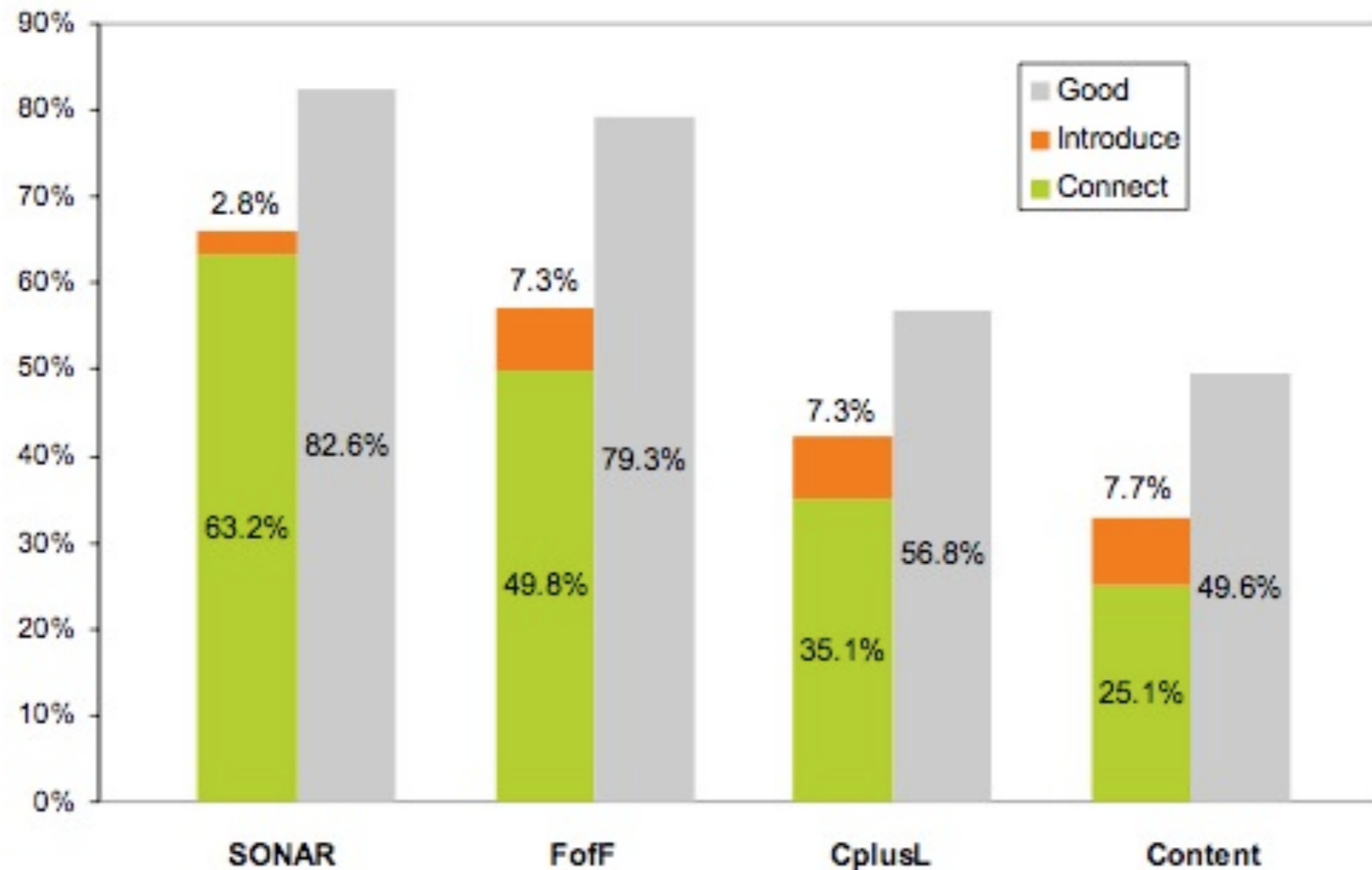
	<b>Content</b>	<b>CplusL</b>	<b>FoF</b>	<b>SONAR</b>
<b>Content</b>		52.8%	1.8%	8.3%
<b>CplusL</b>			3.3%	9.6%
<b>FoF</b>				13.1%

**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.**

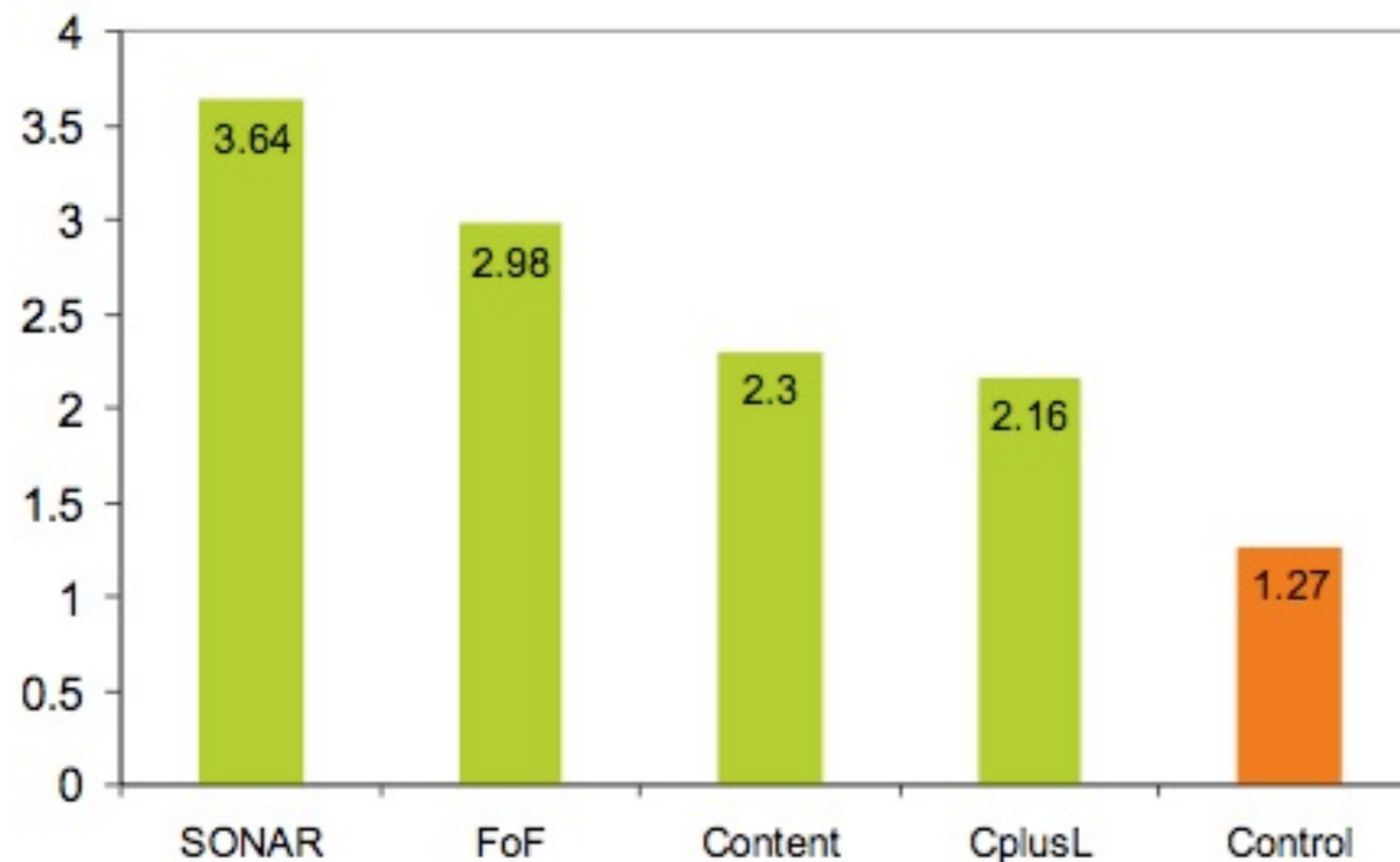


# “Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

[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 interface with the 'Look who else is here' section. The top navigation bar includes 'Home', 'Profile', 'Find People', 'Settings', 'Help', and 'Sign out'. The main heading is 'Look who else is here. Start following them!'. Below this are four tabs: '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, a vertical list of categories is shown, with 'Entertainment' highlighted. The main content area is titled 'Sources in Entertainment' and lists five accounts: MythBusters Official, fearne cotton, Jim Carrey (Verified), Rotten Tomatoes, and Teller (Verified). Each account entry includes a profile picture, name, handle, location, bio, and a 'follow' button.

twitter

Home Profile Find People Settings Help Sign out






## Look who else is here. Start following them!

Browse Suggestions Find Friends Invite By Email Find On Twitter

Look who else is here! Follow the ones you like.

- Art & Design
- Books
- Business
- Charity
- Entertainment**
- Family
- Fashion
- Food & Drink
- Funny
- Health
- Music
- News
- Politics
- Science
- Sports
- Staff Picks
- Staff Picks for Haiti
- Technology
- Travel
- Twitter

### Sources in Entertainment

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



# Twitter-powered Recommendation

The screenshot shows the 'tweetmeme v2' website, which is described as 'Hottest Links on Twitter'. The navigation bar includes categories like Home, Channels, Comedy, Entertainment, Gaming, Lifestyle, Science, Sports, Technology, and World & Bu. Below the navigation bar, there are filters for 'All', 'News', 'Images', and 'Videos'. The main content area is titled 'Everything' and offers sorting options: 'Most Recent', 'Top in 24 Hours', and 'Top in 7 Days'. Three tweets are 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 'Sponsored By Buick' label and a BGR BUICK logo.





# Twitter-powered Recommendation

The screenshot displays the TagWalk website, which provides a "sneaky peek into twitter". The interface includes a search bar, social sharing buttons, and several recommendation widgets:

- 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 grid of images, including a large image of a parliament chamber.
- 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.
- Who's Talking?:** Users talking in English: LuvOrHate weqx techwatching delicious50 EarthTimesPR felloff work\_freelance headlinenews RSSFeedBot Dogbook twinfluence fresh\_projects bananafancy core\_APPLER beafreelancer TechRSS techwatching\_cl mayankchandak IQHQ 4chanbot ZnaTrainer
- 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. Debilitated 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

	<b>BarackObama</b> Location: Washington, DC Bio: 44th President of the United States Similar to: <a href="#">Veronica</a> , <a href="#">katrina_</a> , <a href="#">Pogue</a> <a href="#">See more users like BarackObama</a>	<a href="#">Follow</a>
	<b>Jason</b> Location: Los Angeles, CA Bio: I'm a cereal entrepreneur: Founder of Weblogs, Inc., TechCrunch50, Silicon Alley Reporter, Engadget & Mahalo.com Similar to: <a href="#">Veronica</a> , <a href="#">Scobleizer</a> , <a href="#">TechCrunch</a> <a href="#">See more users like Jason</a>	<a href="#">Follow</a>



# 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

