Introduction to Social Recommendation

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

Department of Computer Science and Engineering
The Chinese University of Hong Kong
Interdependence is and ought to be as much the ideal of man as self-sufficiency.

Man is a social being.

Mahatma Gandhi
A Brief History of the World
A Brief History of the World

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1750</td>
<td>Industrial Revolution</td>
</tr>
<tr>
<td>1945</td>
<td>Information Age</td>
</tr>
<tr>
<td>1969</td>
<td>Birth of Internet</td>
</tr>
<tr>
<td>1975</td>
<td>ENIAC</td>
</tr>
<tr>
<td>1981</td>
<td>IBM Desktop PC</td>
</tr>
<tr>
<td>1983</td>
<td>The MITS Altair</td>
</tr>
<tr>
<td>1984</td>
<td>Apple II</td>
</tr>
<tr>
<td>1989</td>
<td>Time Magazine Person of the Year</td>
</tr>
<tr>
<td>1996</td>
<td>Birth of WWW</td>
</tr>
<tr>
<td>2004</td>
<td>Birth of Web 2.0</td>
</tr>
<tr>
<td>2006</td>
<td></td>
</tr>
</tbody>
</table>

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Top 10 Populations by Countries

as of July 2009

China: 1,335
India: 1,177
United States: 308
Indonesia: 231
Brazil: 192
Pakistan: 168
Bangladesh: 162
Nigeria: 154
Russia: 141
Japan: 127

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Facebook’s Global Audience

Global Audience: 316,402,840

United States
Country Audience: 94,748,820
Percent of Global Audience: 29.95%

United States Male / Female

United States Age Distribution

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
# Facebook’s Growth Stats

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

<table>
<thead>
<tr>
<th>10 Largest Countries</th>
<th>10 Fastest Growing Over Past Week</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. United States</td>
<td>1. Poland</td>
</tr>
<tr>
<td>2. United Kingdom</td>
<td>2. Thailand</td>
</tr>
<tr>
<td>3. Turkey</td>
<td>3. Portugal</td>
</tr>
<tr>
<td>4. France</td>
<td>4. South Africa</td>
</tr>
<tr>
<td>5. Canada</td>
<td>5. Taiwan</td>
</tr>
<tr>
<td>6. Italy</td>
<td>6. Romania</td>
</tr>
<tr>
<td>7. Indonesia</td>
<td>7. Germany</td>
</tr>
<tr>
<td>8. Spain</td>
<td>8. Malaysia</td>
</tr>
<tr>
<td>9. Australia</td>
<td>9. Indonesia</td>
</tr>
<tr>
<td>10. Philippines</td>
<td>10. Iraq</td>
</tr>
<tr>
<td></td>
<td>12.46 %</td>
</tr>
<tr>
<td></td>
<td>137,900</td>
</tr>
<tr>
<td></td>
<td>10.96 %</td>
</tr>
<tr>
<td></td>
<td>161,300</td>
</tr>
<tr>
<td></td>
<td>9.81 %</td>
</tr>
<tr>
<td></td>
<td>80,040</td>
</tr>
<tr>
<td></td>
<td>9.25 %</td>
</tr>
<tr>
<td></td>
<td>189,080</td>
</tr>
<tr>
<td></td>
<td>7.82 %</td>
</tr>
<tr>
<td></td>
<td>367,400</td>
</tr>
<tr>
<td></td>
<td>7.65 %</td>
</tr>
<tr>
<td></td>
<td>28,060</td>
</tr>
<tr>
<td></td>
<td>7.54 %</td>
</tr>
<tr>
<td></td>
<td>350,240</td>
</tr>
<tr>
<td></td>
<td>7.43 %</td>
</tr>
<tr>
<td></td>
<td>236,840</td>
</tr>
<tr>
<td></td>
<td>6.84 %</td>
</tr>
<tr>
<td></td>
<td>752,640</td>
</tr>
<tr>
<td></td>
<td>6.72 %</td>
</tr>
<tr>
<td></td>
<td>6,380</td>
</tr>
<tr>
<td>Alexa as of May 2009</td>
<td>China</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>1</td>
<td>Baidu</td>
</tr>
<tr>
<td>2</td>
<td>QQ</td>
</tr>
<tr>
<td>3</td>
<td>Sina</td>
</tr>
<tr>
<td>4</td>
<td>Google.cn</td>
</tr>
<tr>
<td>5</td>
<td>Taobao</td>
</tr>
<tr>
<td>6</td>
<td>163</td>
</tr>
<tr>
<td>7</td>
<td>Google</td>
</tr>
<tr>
<td>8</td>
<td>Sohu</td>
</tr>
<tr>
<td>9</td>
<td>Youku</td>
</tr>
<tr>
<td>10</td>
<td>Yahoo</td>
</tr>
</tbody>
</table>

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

unfriend
tweet
sexting
defriend
blogsphere
avatar
hashtags
tag cloud
mash-up
twitterati
Folksonomy
blogger
wiki
AVATAR
xjli
nbtivq
cmphfs
Twitter in Spotlight

China’s Great Firewall Blocks Twitter
By ROBERT MACKEY

The Lede
The New York Times News Blog
June 2, 2009, 7:05 PM

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

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

Authority vs. Importance

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
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](#).
Politics

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

<table>
<thead>
<tr>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>2%</td>
</tr>
<tr>
<td>4%</td>
</tr>
<tr>
<td>8%</td>
</tr>
</tbody>
</table>
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!
The YouTube Generation

![YouTube Awards Ceremony Image]
The Age of FaceBook
Social Networking Sites

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

- Social Search Engine
- Leveraging your social networks for searching
Social Media
Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA

Social News/Mash Up

- Digg
- Twitter
- FoxyTunes

What is Twitter?

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: What are you doing?
Social Knowledge Sharing
Social/Human Computation

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

Ling-Jyh Chen, Institute of Information Science, Academia Sinica, Taipei, Taiwan
Human Computation
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
Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Social Recommendations

Genius Recommendations for Apps
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
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.

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Web 2.0 Revolution

- **Glocalization** - think globally and act locally!
- **Weblication** - Web is the application!
- **Three C’s**
  - Connectivity
  - Collaboration
  - Communities
Social Relations

crew
teams
squad
organizations
cohorts
communities
populations
markets
partners
groups

binary
cardinal
integer
real

presence
identity
social role
reputation
expertise
trust
ownership
accountability
knowledge

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

Collective Intelligence
Social Behavior

Intelligent Computation

- ranking
- tagging
- collaborative filtering
- social marketing
- human computation
- opinion mining/sentiment analysis
- query logs analysis
- large graph algorithms
- security & privacy
- NLP
- Regression
- Algorithms
- Model Selection
- Clustering
- Theory
- Classification

- blogs
- wikis
- emails
- instant messaging
- mobile devices
- social bookmarking
- social network services

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

- 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

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


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](https://agora.cs.illinois.edu/display/cs512/home).


References


References


References

• Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, Muthuramakrishnan Venkitasubramaniam, L-diversity: Privacy beyond k-anonymity, TKDD, 2007

• Ninghui Li, Tiancheng Li, and Suresh Venkatasubramaniam, t-Closeness: Privacy Beyond k-Anonymity and l-Diversity, ICDE, 2007.


• 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

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

- Facebook
- Twitter
- LinkedIn
- Google Talk
- Gmail

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
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...**
goglified.com/google-maps-ad-on-chicago-bus/
More results from haochi »

*Google Student Blog: The Google Apps Bus stops at the beginning*
Google Students - connected via twitter.com

Almost two years later, the **Google App to School bus** pulled into Arizona State University and met with over a thousand students, faculty, and staff using ...
googleforstudents.blogspot.com/2008/09/google-apps-bus-stops-at-beginning.html
More results from Google Students »

Searches related to: **google bus**
tamil nadu bus google apps bus google bus routes google bus transit

Results from your social circle for **seattle** - BETA - My social circle - My social content
1078 photos - 17 contacts - Last photo 3 months ago

Results from people in your social circle for **san francisco international airport hotel** - BETA - My social circle - My social content
San Francisco Airport Hotel Burlingame California
Crowne Plaza SFO - connected via twitter.com
Our Burlingame hotel is only 1.5 miles south of San Francisco International Airport on the San Francisco Bay close to an array of exciting attractions. ...
www.sfocp.com/
More results from Crowne Plaza SFO »
Google’s Social Search

News results for jesus
Archbishop of Wales gives his Easter sermon at Llandaff Cathedral - 2 hours ago
*But the Easter story reminds us constantly that God, through Jesus ... She said: "If I were to ask people on the street today 'Have you seen Jesus Christ? ..."
WalesOnline - 1961 related articles »
Taking Up the Dr. Seuss School of Catholicism - TIME - 96 related articles »
Disturbing questions at Easter - Jamaica Gleaner - 93 related articles »

Latest results for jesus - Pause
Jer: It's gonna be 79 today!? Matt: Jesus?
happyinc77 - Twitter - seconds ago
RT @alaintha: @kirstiealley happy jesus resurection day
tinytott67 - Twitter - seconds ago
Jesus Christ Noel, dial down the mental would you? It's Deal or No Deal, not Twin Peaks
doubleshiny - Twitter - seconds ago

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

1. Send Aardvark a question
2. Aardvark finds the perfect person to answer
3. Get their response in a few minutes

What's a great biking path around Golden Gate Park?
Sent 11:36 AM PDT

My favorite is a secret trail that takes you to the beach...
Sent 11:20 AM PDT
Evolution of Search

- Question
  - Contents
  - Machine Intelligence (Dialog systems)
- People
- Friends
- Hybrid
The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]
The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]

- Main components
  - Crawler and Indexer
  - Query Analyzer
  - Ranking Function
  - UI
The Anatomy of A Large-Scale Social Search Engine

-D. Horowitz et al., WWW2010-

• The model
  • With the topics $T$, the probability that user $i$ will successfully answer question $q$ is defined as
    \[
    p(u_i|q) = \sum_{t \in T} p(u_i|t)p(t|q)
    \]
  • Given a question $q$ from user $j$, return a ranked list of user $i$ that maximizes
    \[
    s(u_i, u_j, q) = p(u_i|u_j) \cdot p(u_i|q) = p(u_i|u_j) \sum_{t \in T} p(u_i|t)p(t|q)
    \]
The Anatomy of A Large-Scale Social Search Engine

[D. Horowitz et al., WWW2010]

Figure 3: Example of Aardvark interacting with an answerer

Figure 4: Screenshot of Aardvark Answering Tab on iPhone

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


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?
Information Overload
Real Life Examples

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

Frequently Bought Together
Customers buy this book with Social Network Analysis: A Handbook by John P Scott

Customers Who Bought This Item Also Bought

Five scales rating
I hate it
I don’t like it
It’s ok
I like it
I love it

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

Today's Recommendations For You

Here's a daily sample of items recommended for you. Click here to see all recommendations.

- Invincible ~ Michael Jackson 4 stars (880) $7.99
- In Search of Sunrise, Vol. 7: Asia ~ DJ Tiesto 4 stars (53) $15.99
- Fallen ~ Evanescence 4 stars (2,447) $8.99
- Amar Es Combatir ~ Maná 4 stars (35) $8.49
### Real Life Examples

#### Yahoo! Movies

**My Movies: gabe_ma**  
Edit Profile  

<table>
<thead>
<tr>
<th>Recommendations For You</th>
<th>Receive Recommendations by Email</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Movies in Theaters: 94089</strong></td>
<td></td>
</tr>
</tbody>
</table>

**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!]

**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!]

**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!]

**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!]

**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!]

**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

Songs from friends and similar people

- 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.
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

Q1: How to measure the similarity?

Q2: How to select neighbors?
### User-based Collaborative Filtering

<table>
<thead>
<tr>
<th>Items</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>u3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>u5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u6</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

**Users**

- u1
- u2
- u3
- u4
- u5
- u6

- u1: No recommendations
- u2: Recommendations for items 3 and 4
- u3: Recommendations for items 3 and 4
- u4: Recommendations for items 3, 4, 3, and 4
- u5: No recommendations
- u6: Recommendations for items 3 and 5

# User-based Collaborative Filtering

![Matrix Diagram]

<table>
<thead>
<tr>
<th>Users</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td></td>
</tr>
<tr>
<td>u2</td>
<td>1 3 4 2 5 3 4</td>
</tr>
<tr>
<td>u3</td>
<td></td>
</tr>
<tr>
<td>u4</td>
<td>3 4 3 4 3 4 4</td>
</tr>
<tr>
<td>u5</td>
<td></td>
</tr>
<tr>
<td>u6</td>
<td>1 3 5 2 4 1 3</td>
</tr>
</tbody>
</table>

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
### User-based Collaborative Filtering

<table>
<thead>
<tr>
<th>Users</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td></td>
</tr>
<tr>
<td>u2</td>
<td>1 3 4 2 5 3 4</td>
</tr>
<tr>
<td>u3</td>
<td></td>
</tr>
<tr>
<td>u4</td>
<td>3 4 3 4 3 4 4</td>
</tr>
<tr>
<td>u5</td>
<td></td>
</tr>
<tr>
<td>u6</td>
<td>1 3 5 2 4 1 3</td>
</tr>
</tbody>
</table>

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
User-based Collaborative Filtering

Users

<table>
<thead>
<tr>
<th></th>
<th>u1</th>
<th>u2</th>
<th>u3</th>
<th>u4</th>
<th>u5</th>
<th>u6</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td></td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>u2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>u3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>u4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>u5</td>
<td></td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>u6</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
## User-based Collaborative Filtering

### Users

<table>
<thead>
<tr>
<th>Users</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>u&lt;sub&gt;1&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>u&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1</td>
</tr>
<tr>
<td>u&lt;sub&gt;3&lt;/sub&gt;</td>
<td>3</td>
</tr>
<tr>
<td>u&lt;sub&gt;4&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>u&lt;sub&gt;5&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>u&lt;sub&gt;6&lt;/sub&gt;</td>
<td>1</td>
</tr>
</tbody>
</table>
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

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

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

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

- \( u \): users rated both items
Collaborative Filtering

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

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

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

$$
\begin{pmatrix}
    u_{11} & \cdots & u_{1k} \\
    \vdots & \ddots & \vdots \\
    u_{m1} & \cdots & u_{mk}
\end{pmatrix}
\begin{pmatrix}
    v_{11} & \cdots & v_{1n} \\
    \vdots & \ddots & \vdots \\
    v_{k1} & \cdots & v_{kn}
\end{pmatrix}
$$

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

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

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:

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

- PMF

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

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

\[
p(V | \sigma_V^2) = \prod_{j=1}^{n} \mathcal{N}(V_j | 0, \sigma_V^2 I)
\]
Probabilistic Matrix Factorization

- PMF
  - Bayesian inference

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

My Movie Ratings

<table>
<thead>
<tr>
<th>Movie Title</th>
<th>Rating</th>
<th>Duration</th>
<th>User Rating</th>
<th>Total Ratings</th>
<th>My Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Pursuit of Happyness (PG-13, 1 hr. 57 min.)</td>
<td>A+</td>
<td></td>
<td>B+</td>
<td>38992 ratings</td>
<td>A+</td>
</tr>
<tr>
<td>Finding Nemo (G, 1 hr. 40 min.)</td>
<td>A</td>
<td></td>
<td>B+</td>
<td>137394 ratings</td>
<td>A</td>
</tr>
<tr>
<td>My Blueberry Nights (PG-13, 1 hr. 30 min.)</td>
<td>A+</td>
<td></td>
<td>B-</td>
<td>756 ratings</td>
<td>A+</td>
</tr>
<tr>
<td>Cold Mountain (R, 2 hrs. 35 min.)</td>
<td>B+</td>
<td></td>
<td>B+</td>
<td>38986 ratings</td>
<td>B+</td>
</tr>
<tr>
<td>The Lord of the Rings: The Fellowship of the Ring</td>
<td>A</td>
<td></td>
<td>A</td>
<td>110957 ratings</td>
<td>A</td>
</tr>
<tr>
<td>Shrek 2 (PG, 1 hr. 32 min.)</td>
<td>B</td>
<td></td>
<td>B+</td>
<td>150368 ratings</td>
<td>B</td>
</tr>
</tbody>
</table>
Number of Ratings per User

Extracted From Epinions.com
114,222 users, 754,987 items and 13,385,713 ratings
Challenges

- Traditional recommender systems ignore the social connections between users

Recommendations from friends

Which one should I choose?
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

User-Item Rating Matrix
User-Item Matrix Factorization

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

\[ p(U|\sigma^2_U) = \prod_{i=1}^{m} \mathcal{N}(U_i|0, \sigma^2_U I) \]

\[ p(V|\sigma^2_V) = \prod_{j=1}^{n} \mathcal{N}(V_j|0, \sigma^2_V I) \]

R. Salakhutdinov and A. Mnih \textbf{(NIPS'08)}

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

<table>
<thead>
<tr>
<th></th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$v_5$</th>
<th>$v_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_2$</td>
<td>4</td>
<td></td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_3$</td>
<td></td>
<td>2</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_4$</td>
<td>5</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_5$</td>
<td>5</td>
<td>5</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SoRec

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

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

\[
P(U \mid \sigma_U^2) = \prod_{i=1}^{m} \mathcal{N}(U_i \mid 0, \sigma_U^2 I) \quad P(V \mid \sigma_V^2) = \prod_{j=1}^{n} \mathcal{N}(V_j \mid 0, \sigma_V^2 I)
\]

\[
P(Z \mid \sigma_Z^2) = \prod_{k=1}^{m} \mathcal{N}(Z_k \mid 0, \sigma_Z^2 I)
\]

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

\[
\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}^R (r_{ij} - g(U_i^T V_j))^2 + \frac{\lambda_C}{2} \sum_{i=1}^{m} \sum_{k=1}^{m} I_{ik}^C (c_{ik} - g(U_i^T Z_k))^2
\]

\[
+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2 + \frac{\lambda_Z}{2} \|Z\|_F^2,
\]

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
\[
\frac{\partial L}{\partial U_i} = \sum_{j=1}^{n} I_{ij} g'(U_i^T V_j)(g(U_i^T V_j) - r_{ij})V_j \\
+ \lambda_C \sum_{j=1}^{m} I_{ik} g'(U_i^T Z_k)(g(U_i^T Z_k) - c_{ik}^*)Z_k + \lambda_U U_i,
\]

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

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

• For the Objective Function \( O(\rho_R l + \rho_C l) \)
• For \( \frac{\partial L}{\partial U} \) the complexity is \( O(\rho_R l + \rho_C l) \)
• For \( \frac{\partial L}{\partial V} \) the complexity is \( O(\rho_R l) \)
• For \( \frac{\partial L}{\partial Z} \) 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
Learning to Recommend with Social Trust Ensemble

[Hao Ma, et al., SIGIR2009]
1st Motivation
1st Motivation
1st Motivation

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

2nd Motivation

Ask

Good

Very Good

Cheap & Delicious
2nd Motivation

• Users can be easily influenced by the friends they trust, and prefer their friends’ recommendations.

Where to have dinner?

- Ask
- Good
- Ask
- Very Good
- Ask
- Cheap & Delicious
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

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

\[
p(U|\sigma^2_U) = \prod_{i=1}^{m} \mathcal{N}(U_i | 0, \sigma^2_U I)
\]

\[
p(V|\sigma^2_V) = \prod_{j=1}^{n} \mathcal{N}(V_j | 0, \sigma^2_V I)
\]

[R. Salakhutdinov, et al., NIPS2008]

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

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

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

\[
p(R|S, U, V, \sigma^2_R) = \prod_{i=1}^{m} \prod_{j=1}^{n} \left[ \mathcal{N} \left( R_{ij} | g \left( \sum_{k \in T(i)} S_{ik} U_k^T V_j, \sigma^2_S \right) \right) \right]^{I_{R_{ij}}}
\]
Recommendation with Social Trust Ensemble

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

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
\[ \mathcal{L}(R, S, U, V) \]
\[ = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}^{R}(R_{ij} - g(\alpha U_{i}^{T}V_{j} + (1 - \alpha) \sum_{k \in T(i)} S_{ik}U_{k}^{T}V_{j}))^{2} \]
\[ + \frac{\lambda_{U}}{2} \|U\|_{F}^{2} + \frac{\lambda_{V}}{2} \|V\|_{F}^{2}, \]

\[ \frac{\partial \mathcal{L}}{\partial U_{i}} = \alpha \sum_{j=1}^{n} I_{ij}^{R}g'(\alpha U_{i}^{T}V_{j} + (1 - \alpha) \sum_{k \in T(i)} S_{ik}U_{k}^{T}V_{j})V_{j} \]
\[ \times (g(\alpha U_{i}^{T}V_{j} + (1 - \alpha) \sum_{k \in T(i)} S_{ik}U_{k}^{T}V_{j}) - R_{ij}) \]
\[ + (1 - \alpha) \sum_{p \in B(i)} \sum_{j=1}^{n} I_{pj}^{R}g'(\alpha U_{p}^{T}V_{j} + (1 - \alpha) \sum_{k \in T(p)} S_{pk}U_{k}^{T}V_{j}) \]
\[ \times (g(\alpha U_{p}^{T}V_{j} + (1 - \alpha) \sum_{k \in T(p)} S_{pk}U_{k}^{T}V_{j}) - R_{pj}) \]
\[ \times (\alpha U_{i} + (1 - \alpha) \sum_{k \in T(i)} S_{ik}U_{k}^{T}) + \lambda_{U}U_{i}, \]

\[ \frac{\partial \mathcal{L}}{\partial V_{j}} = \sum_{i=1}^{m} I_{ij}^{R}g'(\alpha U_{i}^{T}V_{j} + (1 - \alpha) \sum_{k \in T(i)} S_{ik}U_{k}^{T}V_{j}) \]
\[ \times (g(\alpha U_{i}^{T}V_{j} + (1 - \alpha) \sum_{k \in T(i)} S_{ik}U_{k}^{T}V_{j}) - R_{ij}) \]
\[ \times (\alpha U_{i} + (1 - \alpha) \sum_{k \in T(i)} S_{ik}U_{k}^{T}) + \lambda_{V}V_{j}, \]
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)

<table>
<thead>
<tr>
<th>Training Data</th>
<th>Metrics</th>
<th>Dimensionality = 5</th>
<th>Dimensionality = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UserMeanItemMean</td>
<td>NMF</td>
<td>PMF</td>
</tr>
<tr>
<td>90%</td>
<td>MAE</td>
<td>0.9134</td>
<td>0.9768</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.1688</td>
<td>1.2375</td>
</tr>
<tr>
<td>80%</td>
<td>MAE</td>
<td>0.9285</td>
<td>0.9913</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>1.1817</td>
<td>1.2584</td>
</tr>
</tbody>
</table>

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

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)

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Impact of Parameter $\alpha$ (Dimensionality = 10)
MAE and RMSE Changes with Iterations

90% as Training Data

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
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>
<thead>
<tr>
<th>Methods</th>
<th>80% Training</th>
<th>50% Training</th>
<th>30% Training</th>
<th>10% Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Mean</td>
<td>0.7686</td>
<td>0.7710</td>
<td>0.7742</td>
<td>0.8234</td>
</tr>
<tr>
<td>Item Mean</td>
<td>0.7379</td>
<td>0.7389</td>
<td>0.7399</td>
<td>0.7484</td>
</tr>
<tr>
<td>SVD</td>
<td>0.6390</td>
<td>0.6547</td>
<td>0.6707</td>
<td>0.7448</td>
</tr>
<tr>
<td>PMF</td>
<td>0.6325</td>
<td>0.6542</td>
<td>0.6698</td>
<td>0.7430</td>
</tr>
<tr>
<td>SoRecUser</td>
<td>0.6209</td>
<td>0.6419</td>
<td>0.6607</td>
<td>0.7040</td>
</tr>
<tr>
<td>SoRecItem</td>
<td><strong>0.6199</strong></td>
<td><strong>0.6407</strong></td>
<td><strong>0.6395</strong></td>
<td><strong>0.7026</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methods</th>
<th>80% Training</th>
<th>50% Training</th>
<th>30% Training</th>
<th>10% Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD</td>
<td>0.6386</td>
<td>0.6534</td>
<td>0.6693</td>
<td>0.7431</td>
</tr>
<tr>
<td>PMF</td>
<td>0.6312</td>
<td>0.6530</td>
<td>0.6683</td>
<td>0.7417</td>
</tr>
<tr>
<td>SoRecUser</td>
<td>0.6197</td>
<td>0.6408</td>
<td>0.6595</td>
<td>0.7028</td>
</tr>
<tr>
<td>SoRecItem</td>
<td><strong>0.6187</strong></td>
<td><strong>0.6395</strong></td>
<td><strong>0.6584</strong></td>
<td><strong>0.7016</strong></td>
</tr>
</tbody>
</table>
Further Discussion of SoRec

- RMSE

<table>
<thead>
<tr>
<th>Methods</th>
<th>80% Training</th>
<th>50% Training</th>
<th>30% Training</th>
<th>10% Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Mean</td>
<td>0.9779</td>
<td>0.9816</td>
<td>0.9869</td>
<td>1.1587</td>
</tr>
<tr>
<td>Item Mean</td>
<td>0.9440</td>
<td>0.9463</td>
<td>0.9505</td>
<td>0.9851</td>
</tr>
<tr>
<td>SVD</td>
<td>0.8327</td>
<td>0.8524</td>
<td>0.8743</td>
<td>0.9892</td>
</tr>
<tr>
<td>PMF</td>
<td>0.8310</td>
<td>0.8582</td>
<td>0.8758</td>
<td>0.9698</td>
</tr>
<tr>
<td>SoRecUser</td>
<td>0.8121</td>
<td>0.8384</td>
<td>0.8604</td>
<td>0.9042</td>
</tr>
<tr>
<td>SoRecItem</td>
<td>0.8112</td>
<td>0.8370</td>
<td>0.8591</td>
<td>0.9033</td>
</tr>
<tr>
<td>SVD</td>
<td>0.8312</td>
<td>0.8509</td>
<td>0.8728</td>
<td>0.9878</td>
</tr>
<tr>
<td>PMF</td>
<td>0.8295</td>
<td>0.8569</td>
<td>0.8743</td>
<td>0.9681</td>
</tr>
<tr>
<td>SoRecUser</td>
<td>0.8110</td>
<td>0.8372</td>
<td>0.8593</td>
<td>0.9034</td>
</tr>
<tr>
<td>SoRecItem</td>
<td>0.8097</td>
<td>0.8359</td>
<td>0.8578</td>
<td>0.9019</td>
</tr>
</tbody>
</table>
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 Ui distrusts user Ud indicates that user Ui disagrees with most of the opinions issued by user Ud.
Distrust

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

\[
\min_{U,V} \mathcal{L}_D(R, S^D, U, V) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(U_i^T V_j))^2
\]

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

\[+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2.\]
Trust

- Users’ trust relations can be interpreted as the “similar” relations
  - On the web, user Ui trusts user Ut indicates that user Ui agrees with most of the opinions issued by user Ut.
Trust

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

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

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

$$+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2.$$
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

<table>
<thead>
<tr>
<th>Statistics</th>
<th>User</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. Num. of Ratings</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Max. Num. of Ratings</td>
<td>162169</td>
<td>1179</td>
</tr>
<tr>
<td>Avg. Num. of Ratings</td>
<td>102.07</td>
<td>17.79</td>
</tr>
</tbody>
</table>

### Table 2: Statistics of Trust Network of Epinions

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Trust per User</th>
<th>Be Trusted per User</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. Num.</td>
<td>2070</td>
<td>3338</td>
</tr>
<tr>
<td>Avg. Num.</td>
<td>5.45</td>
<td>5.45</td>
</tr>
</tbody>
</table>

### Table 3: Statistics of Distrust Network of Epinions

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Distrust per User</th>
<th>Be Distrusted per User</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. Num.</td>
<td>1562</td>
<td>540</td>
</tr>
<tr>
<td>Avg. Num.</td>
<td>0.94</td>
<td>0.94</td>
</tr>
</tbody>
</table>
## Experiments

### RMSE

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Training Data</th>
<th>Dimensionality</th>
<th>PMF</th>
<th>SoRec</th>
<th>RWD</th>
<th>RWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epinions</td>
<td>5%</td>
<td>5D</td>
<td>1.228</td>
<td>1.199</td>
<td>1.186</td>
<td>1.177</td>
</tr>
<tr>
<td></td>
<td>10D</td>
<td></td>
<td>1.214</td>
<td>1.198</td>
<td>1.185</td>
<td>1.176</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>5D</td>
<td>0.990</td>
<td>0.944</td>
<td>0.932</td>
<td>0.924</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10D</td>
<td>0.977</td>
<td>0.941</td>
<td>0.931</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>5D</td>
<td>0.819</td>
<td>0.788</td>
<td>0.723</td>
<td>0.721</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10D</td>
<td>0.818</td>
<td>0.787</td>
<td>0.723</td>
<td>0.720</td>
</tr>
</tbody>
</table>
Impact of Parameters

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


References


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.
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/](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.
Reddit Algorithm

- Time differences
  \[ t_s = A - B \]

- Differences of the up votes and down votes
  \[ x = U - D \]
  \[ y = \begin{cases} 
  1 & \text{if } x > 0 \\
  0 & \text{if } x = 0 \\
  -1 & \text{if } x < 0 
\end{cases} \]
  \[ z = \begin{cases} 
  |x| & \text{if } |x| \geq 1 \\
  1 & \text{if } x < 1 
\end{cases} \]

- Ranking functions
  \[ f(t_s, y, z) = \log_{10} z + \frac{y t_s}{45000} \]

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

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

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 Whathis-face is still a trending topic.
  jam_ie - Twitter - 4 minutes ago
iChat video with front facing camera evidence mounts in iPhone OS ...
  3 Ago 2010 - Fluidstatesnow can explain 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](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 1: Distribution of the Tag Frequency in Flickr.

Figure 2: Distribution of the number of tags per photo in Flickr.

Figure 3: Most frequent WordNet categories for Flickr tags.

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Flickr Tag Recommendation based on Collective Knowledge
[B. Sigurbjörnsson, et al., WWW2008]

Figure 4: System overview of the tag recommendation process.

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
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 Meature:
- Tour Eiffel
- Eiffel
- Seine
- La Tour Eiffel
- Paris

Good at identifying equivalent tags

Asymmetric Meature:
- 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

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

where $P(c \mid u)$ 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
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
      $$\text{interest}^t (\text{category} = c_i) = p^t (\text{click} \mid \text{category} = c_i)$$
      $$= \frac{p^t (\text{category} = c_i \mid \text{click}) p^t (\text{click})}{p^t (\text{category} = c_i)}$$
Personalized News Recommendation Based on Click Behavior

[J. Liu, et al., IUI2008]

• Bayesian Framework for User Interest Prediction

  • Combining predictions of past time periods

\[
\text{interest}(\text{category} = c_i) = \frac{\sum_t \left( N^t \times \text{interest}^t(\text{category} = c_i) \right)}{\sum_t N^t}
\]

\[
= \frac{\sum_t \left( N^t \times \frac{p^t(\text{category} = c_i \mid \text{click}) p^t(\text{click})}{p^t(\text{category} = c_i)} \right)}{\sum_t N^t}
\]

\[N^t\] is the total number of clicks by the user in time period \(t\)

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

\[
\text{interest}(\text{category} = c_i)
\]

\[
p(\text{click}) \times \frac{\sum_t \left( N^t \times \frac{p^t(\text{category} = c_i \mid \text{click})}{p^t(\text{category} = c_i)} \right)}{\sum_t N^t}
\]
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:

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

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

$$p^0(\text{category} = c_i | \text{click}) \propto \frac{\text{interest}(\text{category} = c_i)p^0(\text{category} = c_i)}{p(\text{click})}$$

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

Introduction to Social Recommendation, Irwin King, Michael R. Lyu, and Hao Ma, WWW2010, Raleigh, USA
Personalized News Recommendation Based on Click Behavior

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.

\[
p^0 (\text{category} = c_i \mid \text{click}) = p^0 (\text{category} = c_i) \times \left( \sum_t \left( N^t \times \frac{p^t (\text{category} = c_i \mid \text{click})}{p^t (\text{category} = c_i)} \right) + G \right) \frac{1}{\sum_t N^t + G}
\]
Personalized News Recommendation Based on Click Behavior

• 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

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
User Recommendation

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

Look who else is here. Start following them!

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.)

- fearne cotton
  @FearneCotton
  Location: London
  Bio: Rockin in a free world

- Jim Carrey
  @JimCarrey
  Location: Los Angeles
  Bio: Actor Jim Carrey!

- Rotten Tomatoes
  @RottenTomatoes
  Location: Hollywood, Sydney, London
  Bio: Aggregating reviews from hundreds of movie critics.

- Teller
  @MrTeller
  Location: Bio:
“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 changeling problems
“Make New Friends, but Keep the Old” - Recommending People on Social Networking Sites

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]

<table>
<thead>
<tr>
<th></th>
<th>Content</th>
<th>CplusL</th>
<th>FoF</th>
<th>SONAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td></td>
<td>52.8%</td>
<td>1.8%</td>
<td>8.3%</td>
</tr>
<tr>
<td>CplusL</td>
<td></td>
<td></td>
<td>3.3%</td>
<td>9.6%</td>
</tr>
<tr>
<td>FoF</td>
<td></td>
<td></td>
<td></td>
<td>13.1%</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>SONAR</th>
<th>FoF</th>
<th>CplusL</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
<td>59.7%</td>
<td>47.7%</td>
<td>40.0%</td>
<td>30.5%</td>
</tr>
</tbody>
</table>

Table 2. Recommendations resulting in connect actions.

Figure 4. Increase in number of friends.

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

Look who else is here. Start following them!

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).

fearne cotton
@Fearncotton
Location: London
Bio: rockin in a free world

Jim Carrey
@JimCarrey
Verified
Location: Los Angeles
Bio: Actor Jim Carrey!

Rotten Tomatoes
@RottenTomatoes
Location: Hollywood, Sydney, London
Bio: Aggregating reviews from hundreds of movie critics.

Teller
@MrTeller
Verified
Location: Bio:
Twitter-powered Recommendation

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

• http://blog.oneriot.com/content/2009/06/onерiot-pulse-rank/

• http://uggedal.com/reddit.cf.alorithm.png

• http://www.seopedia.org/tips-tricks/social-media/the-digg-algorithm-unofficial-faq/

• http://www.technologyreview.com/web/24353/?a=f


• 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