Altruism, Selfishness, and Contribution on the Social Web

John Riedl

GroupLens Research
University of Minnesota
Messages

- Web 2.0 is The Social Web
- People Connecting to People
- Technology Enabling Community
1. Google Search
from the frequency that an average surfer uses his or her browser's bookmark feature.

So, the equation is as follows:

\[
P_R(p_i) = \frac{1 - d}{N} + d \sum_{p_j \in M(p_i)} \frac{P_R(p_j)}{L(p_j)}
\]

where \( p_1, p_2, \ldots, p_N \) are the pages under consideration, \( M(p_i) \) is the set of pages that link to \( p_i \), \( L(p_i) \) is the number of links coming from page \( p_i \), and \( N \) is the total number of pages.

The PageRank values are the entries of the dominant eigenvector of the modified adjacency matrix. This makes PageRank a particularly elegant metric: the eigenvector is

\[
R = \begin{bmatrix}
P_R(p_1) \\
P_R(p_2) \\
\vdots \\
P_R(p_N)
\end{bmatrix}
\]

where \( R \) is the solution of the equation

\[
R = \frac{(1 - d)/N}{(1 - d)/N} + d \begin{bmatrix}
\ell(p_1, p_1) & \ell(p_1, p_2) & \cdots & \ell(p_1, p_N) \\
\ell(p_2, p_1) & \ell(p_2, p_2) & \cdots & \ell(p_2, p_N) \\
\vdots & \vdots & \ddots & \vdots \\
\ell(p_N, p_1) & \ell(p_N, p_2) & \cdots & \ell(p_N, p_N)
\end{bmatrix} R
\]

where the adjacency function \( \ell(p_i, p_j) \) is 0 if page \( p_j \) does not link to \( p_i \) and normalized such that, for each \( j \),

\[
\sum_{i=1}^{N} \ell(p_j, p_i) = 1
\]
(Web Search)$^{shared}$

Maurice Coyle and Barry Smyth
AH’08
Want to be a Millionaire?

- Netflix $1M Challenge
Google Street View in UK

- My B&B in London
  - http://maps.google.co.uk/maps?f=q&source=s_q&hl=en&geocode=&q=high+street+kensington&ll=51.490643,-0.158637&sspn=0.010515,0.016243&ie=UTF8&ll=51.505203,-0.193301&spn=0.359.991878&z=17&layer=c&cbll=51.505309,-0.192387&panoid=8ACFOo-EYapAmvgcuRoz_Q&cbp=12,79.63692756901483,0,-8.67291017173075

- Link to BBC Video
  - 0:00 – 2:00

- Privacy Risks
  - Photos of people leaving sex shops
  - Photos of naked toddler playing in park
Opportunity

- How can we mine free activity?
- What are the risks in these data?

2. Yahoo!

Everything
Tag Selection Algorithms

“The Quest for Quality Tags”  
S. Sen, F. Harper, A. LaPitz, J. Riedl  
GROUP 2007
RQ: How can a tagging system show users tags they want to see?

Users don’t agree

Most controversial tags (Bayesian expected entropy):

<table>
<thead>
<tr>
<th>tag</th>
<th>entropy</th>
<th># B</th>
<th># F</th>
</tr>
</thead>
<tbody>
<tr>
<td>comedy</td>
<td>0.987</td>
<td>28</td>
<td>30</td>
</tr>
<tr>
<td>classic</td>
<td>0.986</td>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>stylized</td>
<td>0.983</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>nudity (full frontal)</td>
<td>0.980</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>romance</td>
<td>0.980</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>quirky</td>
<td>0.977</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>magic</td>
<td>0.974</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>animation</td>
<td>0.974</td>
<td>26</td>
<td>20</td>
</tr>
<tr>
<td>Steven Spielberg</td>
<td>0.973</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>sci-fi</td>
<td>0.972</td>
<td>14</td>
<td>17</td>
</tr>
</tbody>
</table>
Tag Prediction

Random baseline: 21%

Implicit features:
- number of applications (39%)
- number of users (51%)
- number of searches for a tag (44%)
- number of users who searched for a tag (48%)
- length of tag (42%)

Moderation-based features:
- global average rating for a tag (59%)
- user-normalized global average rating for a tag (62%)
- tag reputation (57%)

Hybrid combinations: logistic regression, decision trees (67%)
Opportunities

- How can a system distinguish between "good" tags and "bad" tags?
- Can folksonomy be encouraged?
  - Showing users more tags leads to more vocabulary reuse
  - How much convergence is valuable?

3. Facebook

Social Networking for College Students … and everyone else
The Predictive Power of Online Chatter

- Gruhl, Guha, Kumar, Novak, Tomkins
- Yahoo
- ACM KDD 2005

- Volume of blog postings predict sales rank of books
- Queries can be automatically generated in many cases.
- Can sometimes predict spikes in sales rank.

Anti-aliasing on the Web

Jasmine Novak, Prabhakar Raghavan, Andrew Tomkins.
WWW 2004
Story: Finding Medical Records
(Sweeney 2002)

Former Governor of Massachusetts!
- Medical Data
  - Ethnicity
  - Visit Date
  - Diagnosis
  - Procedure
  - Medication
  - Total Charge
- Voter List
  - Name
  - Address
  - Date registered
  - Party affiliation
  - Date last voted

Cat Torturer Video
- YouTube Link
- 1:10 – 2:15
FriendFeed

- Marilyn Katz has updated her profile
- William Ayers has updated his profile
- Antoin "Tony" Rezko is now friends with Nashmi Auichi
- Barack Obama and William Ayers are now friends with Marilyn Katz
- Barack Obama is now friends with William Ayers
- Alexi Giannoulias and Eric Holder have updated their profiles
- Barack Obama is now friends with Alexi Giannoulias
- Barack Obama is now friends with Eric Holder
- Antoin "Tony" Rezko and Allison Davis have updated their profiles
- Barack Obama and Antoin "Tony" Rezko are now friends with Allison Davis
- Barack Obama is now friends with Antoin "Tony" Rezko
Discussion

- Social Implications
- Opportunities
- Threats

4. YouTube

Video by Amateurs?
Copyright issues
- Music videos
- CBS agreement
Why did Google Buy YouTube?

- $1,650 / 65 = $25 million / employee
- $1,650 / 100 million views per day = $16
- $16 / 365 = $.04½ / view / year
- … but Google already had videos!

- The technology?
- The community!

Second Life

- Virtual World
- People “live”, buy, and sell there
- $60M (US $) worth of “manufactured goods” sold this year
Making a Guitar in Second Life

0:45 – 1:45

World of Warcraft

- A different virtual world
- More focus on combat
- 6 million subscribers
Fayejin Funeral

On Tuesday of February 28th Illidan lost not only a good mage, but a good person. For those who knew her, Fayejin was one of the nicest people you could ever meet. On Tuesday she suffered from a stroke and passed away later that night.

5:30 March 4th, Frostfire Hot Springs

World of Warcraft Video
“I hope azshira’s dad dies of a heart attack, then at the funeral some guy runs in naked and pushes the coffin over and runs around slapping people screaming LOL OWNED, then releases a video of it”
WHO'S PLAYING?
A BY-THE-NUMBERS LOOK AT EVERQUEST II'S PLAYERS.

EDUCATION

LESS THAN HIGH SCHOOL: 25.4%
HIGH SCHOOL DIPLOMA: 27.0%
SOME COLLEGE: 28.2%
ASSOCIATES DEGREE: 8.9%
BACHELOR'S DEGREE: 14.6%
GRAD TRAINING OR PROFESSIONAL DEGREE: 6.4%

HEALTH

AVERAGE BMI: ADULT EQI PLAYER 23.9
AMERICAN ADULT 25
CHILDREN AND ADOLESCENT EQI PLAYERS 21.9
CHILDREN AND ADOLESCENT AMERICANS 23.3

PERCENTAGE OF CHEESE ADULT EQI PLAYERS: 37.7%
PERCENTAGE OF CHEESE AMERICAN ADULTS: 30.3%

PERCENTAGE OF MALE EQI PLAYERS DIAGNOSED WITH DEPRESSION: 4.0%
PERCENTAGE OF FEMALE EQI PLAYERS DIAGNOSED WITH DEPRESSION: 5.2%

1M = 1,000,000
Offshore to a Virtual World?

• Nick Yee @ PARC
• Some radiology offshored to India
• Skill in a game: RADAR expert?
  o Learn to detect patterns
  o Rewards for correctness
• Wisdom of Crowds to combine results

Sample Imagery

Example of the size of object to look for. The red rectangle (10px by 4px)  
Example of ocean water.
Chocolate Rain  
by Tay Zonday

- Adam Bahner, a Ph.D. student in American Studies at the University of Minnesota
- Number 2 hottest viral video in history
  - Hottest viral video of Summer 2007
  - Over 26 million views
Fig. 8. Probability of buying a book (DVD) given a number of incoming recommendations.
Maximizing the Spread of Influence through a Social Network, David Kempe, Jon Kleinberg, Éva Tardos, KDD'03

- Independent Cascade Model
  - Information diffuses over time
  - Each neighbor who converts has a one-time chance to convert others

- Linear Threshold Model
  - Each node considers the preferences of all neighbors
  - If total weight passes threshold, a node converts
Video suggestion and discovery for YouTube: Taking random walks through the view graph
Shumeet Baluja, et al., Google, WWW 2008

Opportunities
- Crowd-sourcing
- Gaming as Work?
- How do preferences propagate naturally?
- What predicts fads?
- How do recommenders influence propagation?
Heather Ann Tucci

“I just want to let everyone know August 19 2006 Joe Renner and Joe Shafer died and me and Samatha were hurt. … Both of them knew what they were getting in to. Yes it’s my fault because I was the driver but think about how many of you did what I did.”

Twitter

- Story about VC lunch
- Hudson river pictures
- NYT Tweets during superbowl
- NBA Player who tweeted during halftime of a game
So I once went on a movie date with a guy who thought it was sort of weird that I posted to Twitter about the movie in mid-date. In retrospect, it probably was weird, and a bit rude, and I wouldn't do it again (and no, there was no second date). But get a load of his one.
Discussion

- Social Implications
- Opportunities
- Threats
6. Windows Live
7. MSN

ISP and Content Provider

8. Wikipedia

Next slide, please!
Wikipedia on MySpace

1:20 – 2:15: edit wikipedia to make truth
“What if the number of elephants in Africa were increasing?”
Creating, Destroying, and Restoring Value in Wikipedia
Group 2007

Reid Priedhorsky
Jilin Chen
Shyong (Tony) K. Lam
Katherine Panciera
Loren Terveen
John Riedl
Who contributes Wikipedia’s value?

3.8 million least frequent editors

0.5% of value

14% of value

User: Maveric149

Wales Swartz

PWV contributions of elite editors

Carlson School March 2009
Challenges

- How can vandalism be detected?
- How efficient is Wikipedia?
- How much conflict is valuable?

9. Ebay

Online Auctions
Customers Selling to Customers
Amazon (# 13)

Most Important Resource is Customers
Customers “Selling to” Customers
Today's Hot Trends (USA)

1. scientology is a cult
2. team blackout
3. rod carew
4. you google

4Chan vs. eBaumsWorld

- 4Chan
  - Google Trends Hack
  - Chocolate Rain
- eBaumsWorld
  - Many other hacks
  - “copyright” fight with 4chan
The Internet is Serious Business

“A phrase used to remind those who voluntarily leave the house that being mocked on the Internet is, in fact, the end of the world.”

- Encyclopedia Dramatica
The Social Cost of Cheap Pseudonyms

The Information Cost of Manipulation-Resistance in Recommender Systems
Resnick and Sami. ACM RecSys 08.

Increasing Contributions
**What Theory Tells Us…**

- **Collective Effort Model**
  - People will contribute more if:
    - They believe their effort is important to the group
    - They like the group

- **Smaller is Better**
  - Slovic, Fischhoff, & Lichtenstein, 1980
  - People feel greater concern when the reference group they’re part of grows smaller.

- **Specificity Matters**
  - Small & Loewenstein, 2003
  - Specific identity of those helped is important in drawing people’s support.

---

**CommunityLab Research**

- Social science to increase contributions
  - Accessible to designers
  - Algorithms, interfaces, toolkits

- **GroupLens @ Minnesota**
  - Recommender algorithms and interfaces
  - John Riedl, Joe Konstan, Loren Terveen

- **Bob Kraut and Sara Kiesler @ CMU**
  - Social psychology of computer use

- **Paul Resnick and Yan Chen @ Michigan**
**Voice:** Screen shot

Numerical values are represented by smilies

- ☹️ = Most Valuable
- 😊 = Very Valuable
- 😊 = Valuable

Who the contribution helps

Value of each contribution

---

**Results**

**Self-report**

<table>
<thead>
<tr>
<th>Group</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>3.87</td>
</tr>
<tr>
<td>All MovieLens</td>
<td>3.13</td>
</tr>
<tr>
<td>Similar Group</td>
<td>2.97</td>
</tr>
<tr>
<td>Dissimilar Group</td>
<td>2.94</td>
</tr>
<tr>
<td>Control</td>
<td>2.68</td>
</tr>
</tbody>
</table>

1: Strongly Disagree
2: Disagree
3: Neutral
4: Agree
5: Strongly Agree

---

**Behavioral data**

<table>
<thead>
<tr>
<th>Group</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>7.2%</td>
</tr>
<tr>
<td>All MovieLens</td>
<td>10.2%</td>
</tr>
<tr>
<td>Similar Group</td>
<td>15.8%</td>
</tr>
<tr>
<td>Dissimilar Group</td>
<td>5.9%</td>
</tr>
<tr>
<td>Control</td>
<td>7.4%</td>
</tr>
</tbody>
</table>

---

Want *Smilies* on the regular interface?
Opportunities

- How can contributors be motivated?
- How can social attacks be mitigated?
  - Mail list “unsubscribe”
- How does social psychology interact with defense algorithms?
  - Can the griefers be encouraged to give up?
- Can freedoms be preserved?

10. Craigslist.org

Renting Apartments in NYC
$10/posting: $2.5M/year
Could generate $500M/year with ads
“users haven’t asked for banner ads”
I don't understand—
I've watered down content,
redesigned what was left
and gave it all away for
free on the Internet and
yet I still can't get anybody
to read a newspaper...

Metro Newspaper Publishers

http://www.metaflintrigger.com/Piggy/2003

UNIVERSITY OF MINNESOTA

paper cuts

2007 total: 2,166+ *

* May to December 2007
2009 layoffs / 2008 layoffs / Report buyouts/layoffs

MAP KEY:  1-24  25-49  50-74  75-99  100+  Unknown
paper cuts

2008 total: 15,724+ jobs

UNIVERSITY OF MINNESOTA
Reuters 2nd Life Bureau
Carlson School March 2009

UK to wait and see on virtual world taxes
By PAUL V. PLESS, 29 March 2009 11:48 PM PST

British tax authorities are leaning a closer eye on virtual economies but are
waiting to see the result of US Congressional investigation before taking action.

- Under move by US Congress
- Fine like seen briefly
- Second US bill due to arrive
- UK now more into Second Life’s business
- Tax offices handling for virtual entities

World view Adam Pless

Adam Pless is Reuters’ bureau chief in Second Life.

In real life, he is Adam Poole, a veteran tech and media journalist
He is building regular hours in Second Life, even when it’s a virtual one.

Email him at adam.reuters@gmail.com

http://secondlife.reuters.com
Messages

- Web 2.0 is The Social Web
- People Connecting to People
- Technology Enabling Community

Altruism, Selfishness, and Contribution on the Social Web

John Riedl

GroupLens Research
University of Minnesota
Discussion Topics

- What will happen with virtual economies?
- Why did Google buy YouTube?
- Broadcast -> Narrowcast -> Virtual Life
- Copyright in the Digital Age
- Governing Online Communities
The Social Cost of Cheap Pseudonyms

The Value of Reputation on eBay: A Controlled Experiment

- Web link structure in hyperbolic space
- from Tamara Munzner
Videos for this Presentation, for Kevin

- After (7) YouTube
  - Making a Guitar in Second Life
    - 0:45 – 1:45
  - Suzanne Vega Concert in Second Life
    - 1:00 – 1:40
  - World of Warcraft Video
    - 0:00 – 2:05 (entire video)
- After (9) Wikipedia
  - Wikiality on YouTube
    - 1:20 – 2:15: edit wikipedia to make truth

Bit Bucket
YouTube ‘cat torturer’ traced by web detectives

A teenager who posted videos of himself torturing a cat on YouTube has been arrested after being traced by ‘vigilantes’.

By Matthew Moore
Last Updated: 12:32PM GMT 17 Feb 2009

The youth was shown throwing the animal against a wall, hitting it in the face and holding it under a running shower in the gruesome clips.

In one film titled ‘Potty Cat’ he describes the cat, Dutch, as his ‘best friend’. Another youth can be heard laughing in the background.

The teenager is said to have uploaded the films on YouTube after the original clip was removed but, by the time he returned to the site, the clips had been removed.

The teenager then uploaded the film to his own YouTube account with the words ‘Two of the Best’.

The teenager introduced himself only as Tommy and wore a hoodie to hide his identity, but tech-savvy users pieced together information from other social networking sites including Facebook, MySpace and Spys to come up with a name of a likely culprit.

The teenager then uploaded the film to his own YouTube account with the words ‘Two of the Best’.

The teenager introduced himself only as Tommy and wore a hoodie to hide his identity, but tech-savvy users pieced together information from other social networking sites including Facebook, MySpace and Spys to come up with a name of a likely culprit.