Ranking Mechanisms in Twitter-Like Forums

Anish Das Sarma, Atish Das Sarma, Sreenivas Gollapudi, Rina Panigrahy

Yahoo! Research, Georgia Tech., Microsoft Research, Microsoft Research
Ranking – Ubiquitous

- Web Search
- Blogs
- Forums
- Social Networks
- Question/Answers
- Online Shopping
- Movie Ratings

We explore Ranking based on user ratings.
Screen Shot - Digg

- **Let the Egg show you the way - Shell Shocker Alert**
  - 394 diggs
  - newegg.com - Exclusive Shell Shocker - Don't bury...it's updated daily!

- **Pittsburgh Police Charged for Beating HS Honor Student**
  - 154 diggs
  - cnn.com — An arrest in which several punches were thrown has triggered an accusation of brutality against Pittsburgh, Pennsylvania, police from the mother of the 18-year-old honor student who was taken into custody.

- **National Geographic: Stuck Mars Rover About to Die?**
  - 191 diggs
  - news.nationalgeographic.com — Built to rove for 90 days, Spirit has lasted six years on Mars. But now it's stuck and may lose power by May. Even standing still, though, Spirit can do a surprising amount of science, NASA says.

**Top in All Topics**

- **The Most Amazing MW2 Knife Kill (Video)**
- **The cutest picture you will see today**
- **AT&T Insider Reports Losing iPhone Exclusivity 1/27**
- **OT VII Scientology minister charged with murder**
- **What the ####? Google’s Nexus One Censors Curse Words**
- **Conan Saves His Best Ratings For Last Tonight Show**
- **‘Piracy Isn’t Killing Music’ Radiohead’s Guitarist Says**
Screen Shot - YouTube
Hot or Not?
# NYC Forum

## Threads in Forum: Photos and Videos of New York

<table>
<thead>
<tr>
<th>Thread / Thread Starter</th>
<th>Rating</th>
<th>Last Post</th>
<th>Replies</th>
<th>Views</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sticky: Wired New York Group on Flickr</td>
<td><img src="image" alt="Rating" /></td>
<td>December 31st, 2008 06:40 AM by Lwood</td>
<td>22</td>
<td>16,130</td>
</tr>
<tr>
<td>New York Skyline at Night by Evan Joseph</td>
<td><img src="image" alt="Rating" /></td>
<td>January 29th, 2010 10:59 AM by rp23q7</td>
<td>5</td>
<td>299</td>
</tr>
<tr>
<td>Can you identify building to right of Savoy Plaza?</td>
<td><img src="image" alt="Rating" /></td>
<td>January 28th, 2010 11:57 PM by lofter1</td>
<td>1</td>
<td>82</td>
</tr>
<tr>
<td>Relive the NYC of 1924</td>
<td></td>
<td>January 28th, 2010 05:24 AM by Merry</td>
<td>0</td>
<td>117</td>
</tr>
<tr>
<td>Walking in Bedford Park, The Bronx</td>
<td><img src="image" alt="Rating" /></td>
<td>January 24th, 2010 10:54 PM by Merry</td>
<td>17</td>
<td>5,850</td>
</tr>
<tr>
<td>New Year 2010</td>
<td></td>
<td>January 24th, 2010 10:30 PM by tazar</td>
<td>0</td>
<td>97</td>
</tr>
<tr>
<td>Roofs of New York</td>
<td><img src="image" alt="Rating" /></td>
<td>January 24th, 2010 12:54 PM by Derek2k3</td>
<td>28</td>
<td>12,606</td>
</tr>
</tbody>
</table>
Many more
Popular Ranking Methods

- Star Ratings
- Thumbs up-down (or just Thumbs-up) ratings
- Reputation Points
- Comparisons/Votes
Generic Architecture

KEY QUESTION: How do you pick items for review?
What is a Ranking Mechanism

• How to seek reviews
  – Dissociated from display (floating)
  – Along with display (tied)

• Star-based or Comparison-based? How many?

• What order to display? Or which ones to pick?
  – Can bias users to seek review on what you wish
  – Recent ones? (Users won’t see good ones)
  – Good ones? (Rich gets Richer?)
  – Good but time decayed?

• How to update scores

• Incentive to review/vote.
Metrics for evaluating Mechanism

• Quality of Ranking per cost of review
• Cannot afford too much review bandwidth
  – A small fraction of users viewing items actually take the effort to review
• Bounded Feedback Accurate Rankings
  – Preferably no good item goes unnoticed
  – Only constant reviews. Is this even POSSIBLE?
• Latency
  – Forum writers want feedback quickly to keep them motivated
Our Contributions

• What is a good ranking mechanism?
• Study properties and analyze theoretically different ranking mechanisms
  – Thumbs/Stars-based
  – Comparison-based
• System that implements comparison-based ranking
• Simulations and Experiments
Results

• Our theorems informally
  – Under assumptions of score distribution and voting patterns – accurate rankings possible under comparison based ranking with bounded feedback.
  – Parameters where it is impossible for thumbs-based ranking

• Experimental results (synthetic and real) corroborate theory results

• Comparison-based may be superior to star-based
Possible disadvantages of Stars

Rich gets Richer:
Large volume on all sites.

Unnormalized: Scoring Independent of quality of other item.

Current systems:
Not sufficient incentive for users to vote.
Comparison-based ranking

- Feedback sought in the form of comparisons
  - Users shown a pair of items and have to choose which they prefer

- Advantages:
  - Automatic Normalization
  - Algorithm that converges to accurate rankings faster
  - Requires a bounded (constant) feedback bandwidth
  - Theoretically prove and show through system

- Disadvantages
  - Too expensive?
  - Need to choose pairs (users cannot choose what to review)
Where to use comparisons?

- Public forums:
  - Twitter
  - Digg
  - Jokes sites
  - Ranking messages on social networks (status messages on Facebook)
- Ranking blog posts
- Online shopping sites – Ebay, Amazon
- Generic/Personalized movie recommendations – IMDb
- Flickr, Youtube – applicable but perhaps burdensome (comparing videos)?
- Peer reviewing of papers?
ShoutVelocity System

• [http://shoutvelocity.com](http://shoutvelocity.com)
• Like twitter/digg – can post anything.
• Comparison-based ranking mechanism
• Items picked for review are inversely proportional to current rank estimates
  – Compare with similar approach in Thumbs
  – Even on Digg/Twitter, more popular item likely to be seen (and therefore rated) more.
• We solicit reviews.
### All Top Shouts:

**Aged Top Shouts, Top Shouts in Last: 3 Days, 30 Days, 6 Months, 1 Year**

<table>
<thead>
<tr>
<th>Score</th>
<th>Shout</th>
<th>Comments</th>
<th>Wins</th>
<th>Rank</th>
<th>Author</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.7</td>
<td>75% of all statistics are made up on the spot.</td>
<td>9</td>
<td>38/57</td>
<td>1</td>
<td>Atish Das Sarma</td>
<td>387 days ago</td>
</tr>
<tr>
<td>9.6</td>
<td>&quot;The fellow that agrees with everything you say is either a fool or he is getting ready to skin you.&quot; - Kin H...</td>
<td>2</td>
<td>25/33</td>
<td>2</td>
<td>Akriti Agrawal</td>
<td>200 days ago</td>
</tr>
<tr>
<td>9.5</td>
<td>If it's true that we are here to help others, then what exactly are the others here for?</td>
<td>3</td>
<td>27/34</td>
<td>3</td>
<td>Akash Das Sarma</td>
<td>191 days ago</td>
</tr>
<tr>
<td>9.3</td>
<td>And over there we have the Labyrinth guards. One always lies, one always tells the truth, and one stabs people who ask t...</td>
<td>5</td>
<td>25/31</td>
<td>4</td>
<td>Julia</td>
<td>180 days ago</td>
</tr>
<tr>
<td>9.2</td>
<td>A mathematician is someone who thinks x is a number.</td>
<td>1</td>
<td>18/22</td>
<td>5</td>
<td>Aravind</td>
<td>178 days ago</td>
</tr>
<tr>
<td></td>
<td>Management - Place where ineffective ppl are promoted to where they can</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Hi, Please tell us which shout do you prefer

🌟 aqeel shouts: Anytime you have a 50-50 chance of getting something right, there’s a 90% probability you'll get it wrong :(

Nice one.

(491 chars left)

🌟 Sin shouts: Defeat is not when you fall down, its when you refuse to get up.

Absolutely :)

(487 chars left)
Our System

• **Incentives** for providing feedback:
  – If user wants score/ranking for her post, she is required to provide feedback
    • Therefore, incentive to vote
  – If not, can read and post unhindered
  – Voting allowed even without posting

• **Sufficient Bandwidth?**
  – For every new item that requires a vote we have two previous items voted!
  – Two reviews per item – how to choose which two to review next?
Theoretical Model - Scores

• Distribution of Scores
Probabilistic Voting Model

- Thumbs/Comparisons Scoring Mechanism

![Graph showing the relationship between score, probability, and normal random variables.]

Score: \( x \)

Thumbs prob: \( f(x) \)

Comparisons

Prob. of score \( x \) beating score \( y \)

\[ f(x - y) \]

\[ f(x) = \text{erf}(\alpha x) \]

Probability that normal random variable with mean 0 variance \( \alpha/\sqrt{2} \) exceeds \( x \)
Theory Results

• Comparisons: If scores normal distribution,
  – Algorithm to estimate rank $r$ to within $r(1 \pm \epsilon)$
  – Bounded feedback per item.

• Thumbs: If scores normal distribution
  – If $(\alpha > 1)$
  – No algorithm can approximate rankings
Algorithm (Intuition for Static)

• Static Tournament to estimate score of each item.
Algorithm (Dynamic)

- Dynamic tournament using a queue
Update Scores (similar to ELO in chess)

[Arpad Elo, 1978]

Discount factor (to capture intuition that after lot of comparisons, scores stabilize)

\[ c_i = \frac{2}{(1 + k_i)^{0.5}} \]

If \( i_1 \) beats \( i_2 \):

\[ s_1 = s_1 + c_1 \times \left(1 - \frac{1}{1 + 2(s_2 - s_1)}\right) \]

\[ s_2 = s_2 - c_2 \times \left(\frac{1}{1 + 2(s_1 - s_2)}\right) \]
Experimental Evaluation

• Simulations: (synthetic)
  – 1000 players with normal distribution scores (variance 1)
  – Twice the number of thumb evaluations as the number of comparisons

• Experiments on ShoutVelocity: (real)
  – 4853 comparisons over 1245 items
  – True score not known so estimated using many random ELO rated tournaments
Simulation

At 1000, 2000 reviews (comparison, thumbs), MRR are 0.244 and 0.019
Experiments

MRR shoots up fast
ShoutVelocity System Latency

Number of reviews to wait before getting first review
Scores (CDF) on ShoutVelocity

Assumption that Normal distribution is reasonable.
Conclusions

• Comparison Based Ranking presented
  – different categories? can one say unrelated/equal?

• ShoutVelocity System

• Star ratings
  – Generalization of Thumbs (and theoretical analysis can easily be extended)

• Generalize pairwise to n-wise comparisons?
  – Practicality questionable (more significant burden on users)

• Paper reviewing? 😊
Ranking Mechanisms in Twitter-Like Forums

Anish Das Sarma, Atish Das Sarma, Sreenivas Gollapudi, Rina Panigrahy

Yahoo! Research, Georgia Tech., Microsoft Research, Microsoft Research

On the job market!
atish@cc.gatech.edu
Thank You!