Thrust III. Modeling and Experiments

• Data-driven study of biases
  • Modeling biases in user-based sequence data
  • Behavioral biases in group formation

• Behavioral experiments on Online Platforms
  • Experiments on belief formation in groups
  • Effect of native advertising on processing of news and its credibility
Biases/behaviors in sequence data

• User-based sequence data is incredibly common (music, web traversal, location visits, emails)

• Repeat behavior bias [Benson et al. 16, 18] play the same song, visit same web site, watch the same video

• Recency behavior bias [Benson et al. 16, 18] play recent songs and videos

• Possible explanations and models?
Single-item and set-based sequences

Single-item
• music and video plays
• clicks on the web
• Virtual “check-ins”

Set-based
• Email recipients
• Tags on questions on Stack Overflow
• Coauthorship
• In-person group interactions
Repeat behaviors/biases in single-item sequence data

- **last.fm**: 69%
- **YouTube**: 26%
- **Wikipedia**: 15%
- **Google+**: 31%
- **Brightkite**: 51%
- **Clicks**: 38%
- **Check-ins**: 31%
- **Music and video**: 44% sequences of plays
- **Clicks**: navigation on the web
- **Check-ins**: physical places visited
Repeats in set sequence data

![Graphs showing the fraction of exact repeats and superset repeats across set sizes for different datasets.](image-url)
Recency bias in single-item sequence data

Pr(repeat from j steps back) \propto w_j

“The Dynamics of Repeat Consumption”.
A. Anderson et al., 2014.
Recency bias in set sequence data

![Graph showing relative Jaccard index for different datasets over position prior (k).](image-url)

- **tags-mathoverflow**
- **tags-math-sx**
- **email-Enron-core**
- **email-Eu-core**
- **contact-prim-school**
- **contact-high-school**
- **coauth-Business**
- **coauth-Geology**
Repeats and recency in sequence data: bounded rationality in exploration?

- Exploit-explore tradeoff? Not enough *exploration*.
- How to understand this as rational?
  - People have limited exploration capabilities.
  - Value gained from future maybe discounted.
  - People make quick decisions.
  - People have limited memory of utilities
  - Exploration bursts?
Biases in sequence data: satisficers vs. maximizers?

(Barry Schwartz)

- Satisficers: choose something if “good enough” (above threshold)
- Maximizers: choose to maximize utility in some way
- Can who is satisficer and who is maximizer from data?
Biases and behaviors in group-formation

• How do groups form? Evolve? Succeed?

• Search space grows exponentially in size (too much to explore)

Approaches:
1. Set sequences as group formation
2. Network-based analysis of group formation
Correlation in set sequences

- For each sequence in each dataset, we count the number of times each size-2 and size-3 subset appears.
- We then count the same statistics under a null model where elements are randomly placed into sets.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>size-2 subset counts</th>
<th></th>
<th>size-3 subset counts</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>data</td>
<td>null model</td>
<td>data</td>
<td>null model</td>
</tr>
<tr>
<td>email-Enron-core</td>
<td>5.82</td>
<td>4.34 ± 0.043</td>
<td>4.23</td>
<td>2.67 ± 0.038</td>
</tr>
<tr>
<td>email-Eu-core</td>
<td>4.46</td>
<td>3.11 ± 0.008</td>
<td>3.23</td>
<td>2.08 ± 0.007</td>
</tr>
<tr>
<td>contact-prim-school</td>
<td>2.36</td>
<td>1.87 ± 0.003</td>
<td>1.35</td>
<td>1.09 ± 0.002</td>
</tr>
<tr>
<td>contact-high-school</td>
<td>4.49</td>
<td>3.26 ± 0.007</td>
<td>2.09</td>
<td>1.35 ± 0.004</td>
</tr>
<tr>
<td>tags-mathoverflow</td>
<td>1.49</td>
<td>1.41 ± 0.002</td>
<td>1.18</td>
<td>1.15 ± 0.002</td>
</tr>
<tr>
<td>tags-math-sx</td>
<td>1.49</td>
<td>1.31 ± 0.001</td>
<td>1.21</td>
<td>1.12 ± 0.001</td>
</tr>
<tr>
<td>coauth-Business</td>
<td>1.50</td>
<td>1.30 ± 0.001</td>
<td>1.40</td>
<td>1.24 ± 0.001</td>
</tr>
<tr>
<td>coauth-Geology</td>
<td>1.29</td>
<td>1.15 ± 0.000</td>
<td>1.15</td>
<td>1.07 ± 0.000</td>
</tr>
</tbody>
</table>
Our correlated repeated unions model captures repeats, recency, and correlations.

Setup.
Observe sequence of sets $S_1, ..., S_k$.
Given number $r$ of repeated elements in $S_{k+1}$.
Model selects $r$ elements from $\bigcup_{j=1}^{k} S_j$.

CRU model.
Start with $N = \emptyset$, given $r$.
1. Sample set $S_{k-j}$ from $j$ steps back with recency weight $w_j$.
2. Sample $T$ by keeping each item $x$ in $S_{k-j}$ with correlation probability $p$.
3. $N = N \cup T$.
4. Repeat steps 1—3 until $|N| = r$.
   (if $T$ makes $N$ too large, randomly drop elements from $T$)
Set-building in group-formation

• Copying in coauthorship
• On average, an author’s coauthors on a paper sample recent paper co-authors at 80-90% correlation level
• Possible rational model: local search & nearby local optimum

Mean per-set likelihood

CRU model.
Baseline model (flat, no structure). Similar to [Anderson+ 14]
Recency in group-formation (set sequences)

Recency weight $w$

coauth-Geology

Correlation probability $p$. 

<table>
<thead>
<tr>
<th>Probability</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>blue</td>
</tr>
<tr>
<td>0.1</td>
<td>orange</td>
</tr>
<tr>
<td>0.2</td>
<td>green</td>
</tr>
<tr>
<td>0.3</td>
<td>red</td>
</tr>
<tr>
<td>0.4</td>
<td>pink</td>
</tr>
<tr>
<td>0.5</td>
<td>brown</td>
</tr>
<tr>
<td>0.6</td>
<td>purple</td>
</tr>
<tr>
<td>0.7</td>
<td>gray</td>
</tr>
<tr>
<td>0.8</td>
<td>yellow</td>
</tr>
<tr>
<td>0.9</td>
<td>cyan</td>
</tr>
<tr>
<td>0.99</td>
<td>blue</td>
</tr>
</tbody>
</table>
Network-based in group-formation behavior

• High-level features [Benson-Abebe-Schaub-Jadbabaie-Kleinberg 18]

Coauthorship data of scholars publishing in history. $W_{ij} = \#$ of papers containing nodes $i$ and $j$. 
Biases in group-formation (coauthorship)

- High-level features [Benson-Abebe-Schaub-Jadbabaie-Kleinberg 18]
Biases/behaviors in group-formation

What role do various features play in group formation bias?
• Proximity? (authors from same institution / same MURI)
• Fame / popularity?
• Opinion dynamics?

What biases are involved in group evolution / growth / success?
• Growth from proximity?
• Growth from diversification?
• Change biased by context?
Thrust III. Modeling and Experiments

• Data-driven study of biases
  • Modeling biases in user-based sequence data
  • Behavioral biases in group formation

• Behavioral experiments on Online Platforms
  • Experiments on belief formation in groups
  • Effect of native advertising on processing of news and its credibility
Experiments on belief formation in groups

- Experimental identification of belief formation mechanisms in groups

- We design human subject experiments to test and compare plausible models of how people aggregate other beliefs and update their own.
Belief formation in groups

- Belief aggregation among group members who exchange their opinions:
  - Jury deliberations, expert committees, medical diagnoses
  - decision-making organizations, such as
    - legal or medical consulting firms and the public sector bureaus.
  - Non-Bayesian aggregators such as arithmetic or geometric averages are known to be susceptible to repetitions, persuasion, and social confirmation biases (DeMarzo et al, Eyster and Rabin).
  - Our experimental results help identify the theoretical paradigms that describe the aggregate beliefs and their susceptibility to such biases.
Experimental platform

- A prototype of our experimental platform can be found here: http://segregation.media.mit.edu/
- We recruit our subjects from Amazon MTurk.
- In four hypothetical scenarios, we ask them to evaluate
  - a job candidate,
  - a case for insurance,
  - an investment opportunity,
  - a stock position.
- They do so after learning about the nature of uncertainties and being randomly exposed to different experimental conditions that consist of other reported beliefs.

In the next three slides we provide snapshots from the experiment platform.
### Initial impression
- **Startup 1**: ✓
- **Startup 2**: ✓
- **Startup 3**: ✓
- **Startup 4**: 

### Your decision
- **Startup 1**: 
- **Startup 2**: 
- **Startup 3**: 
- **Startup 4**: 

### Startup performance
- **Startup 1**: 
- **Startup 2**: 
- **Startup 3**: 
- **Startup 4**: 

### Feedback
You’re correct! In this case, it was a **bad** decision to invest in the startup, even though the initial report was **positive**.

White papers are never totally reliable indicators of whether it is a good idea to fund a startup, but you have to start somewhere, so you read the startup’s white paper.

The impressions you receive from a white paper can be either favorable or unfavorable.

You receive a **favorable** impression from reading the startup's white paper.

Based just on the information you have received so far, how much do you believe it is a good idea to fund the startup?
<table>
<thead>
<tr>
<th>Startup 1</th>
<th>Startup 2</th>
<th>Startup 3</th>
<th>Startup 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Your decision</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
</tr>
<tr>
<td>Startup performance</td>
<td>▼</td>
<td>✓</td>
<td>▼</td>
</tr>
<tr>
<td>Feedback</td>
<td>You're correct! In this case, it was a <strong>bad</strong> decision to invest in the startup., even though the initial report was <strong>positive</strong>.</td>
<td>Sorry, In this case, it was a <strong>good</strong> decision to invest in the startup.</td>
<td>You're correct! In this case, it was a <strong>bad</strong> decision to invest in the startup.</td>
</tr>
</tbody>
</table>

The goal of these trials was for you to realize that when you have a favorable impression of a case there is still a significant (one in four) chance that the case is bad. Similarly, when you have an unfavorable impression there is still a significant (one in four) chance that the case is good. That is why it is a good idea to rely on other people's opinion to make better-informed decisions.

In the next two rounds, we will ask you to form an opinion of the case, based on what you hear from your colleagues. Your colleagues will help you by reporting their own evaluations of the case.

[Continue]
You have not yet done any research on this startup yourself, but two of your partners in your firm researched the startup and report back to you.

Your partners can be totally certain, fairly certain, somewhat certain, or totally uncertain about it being a good or a bad idea to fund the startup.

Your first partner reports to you that she currently is **fairly certain** that it is a **bad** idea to fund the startup.

Your second partner reports to you that he currently is **fairly certain** that it is a **good** idea to fund the startup.

Based on the information you have received up to now, how much do you believe it is a good idea to fund the startup?
Results

Our preliminary findings suggest that arithmetic and geometric averaging provide a better description of the empirical data, compared to the Bayesian posterior aggregator.
Punchline

- Linear and geometric averaging are better fits to the observed data (compared to the Bayes or no-recall models)

- They often provide similar predictions
  - Geometric averaging has the zero belief contagion property (similar to the Bayes rule)
  - Linear averaging has the positive belief contagion property
  - We can design experimental conditions to identify these models based on the zero and positive belief contagion properties
Most Americans say they have lost trust in the media

Columbia Journalism Review
MOTIVATION

CONTENT RECOMMENDATION NETWORKS (CRNS)
CRNs are an intermediary between publishers and advertisement companies.

CRNs aggregate third-party ads that are embedded on the online newspapers' webpage.

CRNs also recirculate news articles from the host publisher (house ads).
RESULTS

DESCRIPTIVE ANALYSIS: THIRD-PARTY ADS

- The majority of CRN's ads are perceived as **clickbait**.
- CRNs convey dubious **political** information through ads.

Over **80%** of the ads were clickbait.

An average of **11%** of the ads were detected to be political.

Do CRNs undermine political credibility of mainstream news outlets?
METHOD

BEHAVIORAL EXPERIMENT

- Study set-up on Lucid:
  5,000 respondents evaluated the credibility of articles from CNN, FoxNews, The Atlantic and SacramentoBee with and without CRNs

- Tool for analysis:
  Fisherian randomization: use randomisation to make inference about causal effect (under the null hypothesis, for average treatment effect)
RESULTS

BEHAVIORAL EXPERIMENT

- CRNs impact credibility of less well-known publishers The Atlantic and Sacramento Bee.

CRNs decrease trustworthiness of The Atlantic by 2.3% among those familiar with the outlet (2k).

CRNs increase trustworthiness of Sacramento Bee by 1.8% among those unfamiliar with the outlet (4k).

Will redo the survey with more articles and more CRNs
RESULTS

AVERAGE TREATMENT EFFECTS (ON A 5-POINT SCALE) OF OLS REGRESSION ON EACH PUBLISHER.