The Dynamics of Repeat Consumption

Ashton Anderson
Department of Computer Science
Stanford University

Ravi Kumar, Andrew Tomkins, Sergei Vassilvitskii
Google, Inc.


2014.09.30
Yunhyoung Kim
Table of Contents

• Introduction
• Analysis on Datasets
• Consumption Model
• Experiments
• Conclusion
Introduction

Researches on “why people reconsume”

Hedonic Psychology
- utility of consuming an item repeatedly
- Indulgence or Variety-Seeking

Brand Choice
- consumer choice is a stochastic process
- Brand switching occurs

Repeat Queries
- over 40% queries are repeat queries
- refinding information

Website Revisit
- revisit due to content change
- empirical analysis on revisit pattern

Purpose of this paper

Develop a model to explain the repeat behavior of customers
Introduction

Organization of this paper

Collected Data

Analyze Data & Extract Features

Modeling & Fitting

Experiments & Observation

Datasets

Location-based Services

Wikipedia content access

MapClicks

WikiClicks

Yes.com

REWfm

radio playlists from all radio stations in US

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#users</th>
<th>#unique items</th>
<th>frac unique / user</th>
<th>time span</th>
</tr>
</thead>
<tbody>
<tr>
<td>BrightKite</td>
<td>51.4K</td>
<td>773K</td>
<td>0.55</td>
<td>2008–2010</td>
</tr>
<tr>
<td>GPlus</td>
<td>18.4K</td>
<td>1.81M</td>
<td>0.51</td>
<td>2006–2013</td>
</tr>
<tr>
<td>MapClicks</td>
<td>431K</td>
<td>216K</td>
<td>0.62</td>
<td>2006–2013</td>
</tr>
<tr>
<td>Shakespeares</td>
<td>6403</td>
<td>26</td>
<td>0.31</td>
<td>1589–1613</td>
</tr>
<tr>
<td>WikiClicks</td>
<td>852K</td>
<td>529K</td>
<td>0.88</td>
<td>2005–2013</td>
</tr>
<tr>
<td>Yes</td>
<td>15.8K</td>
<td>75.2K</td>
<td>0.79</td>
<td>2010–2011</td>
</tr>
<tr>
<td>YouTube</td>
<td>696K</td>
<td>1.44M</td>
<td>0.83</td>
<td>2011–2013</td>
</tr>
<tr>
<td>WikiClicks-Food</td>
<td>298K</td>
<td>36.9K</td>
<td>0.68</td>
<td>2011–2013</td>
</tr>
<tr>
<td>YouTube-Music</td>
<td>694K</td>
<td>497K</td>
<td>0.78</td>
<td>2011–2013</td>
</tr>
</tbody>
</table>

Baseline (not generated by consumption)
Analysis on Datasets

Complementary CDF of the number of consumption

CCDF of the number of items consumed

Distribution of fraction of repeat consumption
Analysis on Datasets

Popularity

Past performance predicts the future VS. Users are variety seeking

The item a user has already consumed many times is likely to be repeatedly consumed again in the future

Satiation

Assumption: a user will be bored about the repetition of the same consumption

They checked:
Analysis on Datasets

Recency

Recency: tendency for more recently-consumed items to be reconsumed (than items consumed further in the past)

Even though C appears only once, C is more likely to be reconsumed than A

Is this true? Let’s see using cache.

**ABBAEAEAFEBDBEFDBCD**

Optimal Replacement: Replace the furthest item in the future

**ABE**  **FBE**  **DBE**  **DBF**  **DCF**

Compare this with 1) Random Permutation with optimal replacement and 2) Stable top-k cache

**DDBAFCEAEFBBDBEAEAB**

---

(i) **MapClicks**  
(ii) **BrightKite**  
(iii) **Shakespeare**
## Analysis on Datasets

**What we have seen until now**

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popularity</td>
<td>The item a user has already consumed many times is likely to be repeatedly consumed again in the future</td>
</tr>
<tr>
<td>No satiation</td>
<td>An item which is consumed in a row is likely to be consumed in the future</td>
</tr>
<tr>
<td>Recency</td>
<td>Recently-consumed items are more likely to be consumed again in the future (than items consumed further in the past)</td>
</tr>
</tbody>
</table>

Based on these analyses, consumption model is suggested
## Consumption Model

### Popularity vs. Recency

Which one is more influential factor on repeat consumption?

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality model</td>
<td>Quality of an item is the primary factor (Popularity = one aspect of quality)</td>
</tr>
</tbody>
</table>
|                | \[
| score \, s(e) | each item \, e \in E                                                        |
|                | Probability an item \, e is selected : \, \frac{s(e)}{\sum_{e' \in E} s(e')} |
| Recency model  | Recently consumed items are more likely to be consumed                        |
| \(w(i - j)\)  | weight consuming an item seen at time \, j < i                                |
| \(w(1)\)      | weight for repeating the previously consumed item                            |
|                | Probability an item \, e is selected : \, \frac{\sum_{j < i} I(x_i = e) w(i - j)}{\sum_{j < i} w(i - j)} |
| Hybrid model   | Combine quality and recency model                                            |
|                | Probability an item \, e is selected : \, \frac{\sum_{j < i} I(x_j = e) w(i - j) s(x_j)}{\sum_{j < i} w(i - j) s(x_{i-j})} |
Consumption Model

Recency model example

\[ \sum_{j<i} I(x_i = e)w(i - j) \]
\[ \sum_{j<i} w(i - j) \]

Probability that item A is selected:
\[ \frac{w(12) + w(14) + w(16) + w(19)}{w(1) + w(2) + \cdots + w(19)} \]

Probability that item B is selected:
\[ \frac{w(3) + w(7) + w(9) + w(17) + w(18)}{w(1) + w(2) + \cdots + w(19)} \]

\( w(i) \geq w(i + 1) \quad \forall i \) : recent consumption has more weight

Hybrid model example

\[ \sum_{j<i} I(x_j = e)w(i - j)s(x_j) \]
\[ \sum_{j<i} w(i - j)s(x_{i-j}) \]

Probability that item A is selected:
\[ \frac{4/19 \cdot (w(12) + w(14) + w(16) + w(19))}{3/19 \cdot w(1) + 1/19 \cdot w(2) + \cdots + 4/19 \cdot w(19)} \]
Consumption Model

Learning model parameters

Quality model

\[ s(e) = \frac{1}{k} \sum_{i=1}^{k} I(x_i = e) \]

Recency model

applied gradient ascent method

\[
LL = \sum_{i \in R} \log \left( \sum_{j < i} I(x_i = x_j)w(i - j)s(x_j) \right) - \sum_{i \in R} \log \left( \sum_{j < i} w(i - j)s(x_j) \right)
\]

Calculate \( \frac{\partial LL}{\partial w(\delta)} \), \( \frac{\partial LL}{\partial s(e)} \)

\[
\frac{\sum_{j < i} I(x_j = e)w(i - j)s(x_j)}{\sum_{j < i} w(i - j)s(x_{i-j})}
\]

LL is not concave in s and w, it is not guaranteed to find global optimum

**Tipping behavior**

\( w(i) \geq w(i + 1) \quad \forall i \)

If \( \sum_{i=1}^{\infty} w_i \) converges, one item is ever consumed after some time \( \tau \)

If \( \sum_{i=1}^{\infty} w_i \) diverges, every item is consumed infinitely many times
Experiments

Distributions of recency weights

Popularity vs. Recency - which one reflects more?

<table>
<thead>
<tr>
<th></th>
<th>$s(\cdot)$ = popularity</th>
<th>$w(\cdot)$ = learned</th>
<th>uniform</th>
<th>learned</th>
<th>learned</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIGHTKITE</td>
<td>0.375</td>
<td>0.617</td>
<td>0.637</td>
<td>0.936</td>
<td>1</td>
</tr>
<tr>
<td>GPLUS</td>
<td>0.587</td>
<td>0.801</td>
<td>0.794</td>
<td>0.877</td>
<td>1</td>
</tr>
<tr>
<td>MAPCLICKS</td>
<td>0.383</td>
<td>0.931</td>
<td>0.414</td>
<td>0.989</td>
<td>1</td>
</tr>
<tr>
<td>WIKICLICKS</td>
<td>0.503</td>
<td>0.724</td>
<td>0.687</td>
<td>0.945</td>
<td>1</td>
</tr>
<tr>
<td>YOUTUBE</td>
<td>0.636</td>
<td>0.677</td>
<td>0.924</td>
<td>0.962</td>
<td>1</td>
</tr>
</tbody>
</table>
Experiments

Fit to a Power Law with Exponential CutOff (PLECO)

PLECO is found in many natural and man-made phenomena

\[ \Pr(x) \propto (x + \gamma)^{-\alpha} e^{-\beta x} \]

<table>
<thead>
<tr>
<th>Dataset</th>
<th>(\alpha)</th>
<th>(1/\beta)</th>
<th>(\gamma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIGHTKITE</td>
<td>1.8</td>
<td>670</td>
<td>10</td>
</tr>
<tr>
<td>MAPCLICKS</td>
<td>1.42</td>
<td>1000</td>
<td>10</td>
</tr>
<tr>
<td>GPLUS</td>
<td>1.95</td>
<td>1250</td>
<td>7</td>
</tr>
<tr>
<td>SHAKESPEARE</td>
<td>0.15</td>
<td>83</td>
<td>10</td>
</tr>
<tr>
<td>WIKICLICKS</td>
<td>1.01</td>
<td>2222</td>
<td>5</td>
</tr>
<tr>
<td>YOUTUBE</td>
<td>1.16</td>
<td>2000</td>
<td>10</td>
</tr>
</tbody>
</table>
Experiments

Does additivity of recency weights holds empirically?

$$\frac{\sum_{j<i} I(x_i = e)w(i-j)}{\sum_{j<i} w(i-j)}$$  Probability an item is selected is proportional to:  \( w(i) + w(j) \)

What about \( w(i, j) \)?

Calculate \( w(i, j) \) from empirical data, and compare it with \( w(i) + w(j) \)

\[
\log \left( \frac{w(i, j)}{w(i) + w(j)} \right) \quad \hat{w}(i, j) \quad \text{excluding the first appeared items}
\]
Conclusion

• Suggest an additive model of recency
• Show that additivity is good fit to data
• Inferred values are well-fit by PLECO
• Propose hybrid model (recency+quality) which outperforms models based only on quality or recency