Improving News Personalization through Search Logs*

Xiao Bai
Yahoo Research, USA
xbai@oath.com

B. Barla Cambazoglu
Independent Researcher, Spain
barla@berkantbarlacambazoglu.com

Francesco Gullo
UniCredit, R&D Dept., Italy
gullof@acm.org

Amin Mantrach
Criteo Research, USA
a.mantrach@criteo.com

Fabrizio Silvestri
Facebook, UK
fsilvestri@fb.com

* The work was carried out when all the authors were affiliated with Yahoo Labs, Barcelona, Spain
Motivation

• Personalization of a news service is a long-standing challenge
  • Traditional approaches: ranking news articles based on how well they match the user profile

• Existing user profiles are built using endogenous information only

• Exogenous information also matters
Motivation

Why does *exogenous information* matter?

- A user from Europe accesses the news service for football-related news
- The same user is planning a trip to the US and starts querying some search engine about flights and accommodation in the US
- Assume significant changes in the rules for European citizens to enter the US become public
- Such changes would clearly be a news of interest to the user, but it would not be recommended if the user profile was built by considering only endogenous (football-related) information
Contributions

We study the novel problem of news personalization by leveraging web-search logs

- Understand what kind of information in search logs should be considered to build more complete user profiles
- Methods for constructing and combining user profiles are beyond the scope of the work
- Thorough experimental evaluation to answer 5 critical research questions
Methodology: Constructing user profiles

• We focus on users who have both used the online news service and the web-search service
  • We build *news profiles* and *search profiles*
  • Computed as TF-IDF vectors of the terms in the news read in the past (news profiles) or queries issued to the search engine (search profiles)

• 3 granularities of search profile:
  • Query-based
  • Title-enriched
  • Abstract-enriched
Methodology: Combining user profiles

• 2 scores:
  • News scores (cosine similarity between news profile and news vector)
  • Search score (cosine similarity between search profile and news vector)

• 2 methods:
  • Score aggregation: SP_Score method
  • Rank aggregation: SP_Rank method
Experiments - Setting

- Dataset: click logs of Yahoo News and query logs of Yahoo Web Search
- We picked a random day
- Queries issued up to 6 months before the picked day
- Users issued at least 1000 queries in the 3 months before the picked day
- 70K users, 140K independent news recommendations yielded in the picked day
- Methods: proposed SP_Score and SP_Rank vs. a baseline B that relies on the news profiles only
- Performance assessment: NDCG metric
Research Question 1: Do search profiles improve the quality of news personalization?

<table>
<thead>
<tr>
<th></th>
<th>avg. NDCG</th>
<th>p-value&lt;0.05 (vs. B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.522</td>
<td>—</td>
</tr>
<tr>
<td>SP_Score</td>
<td>0.545</td>
<td>yes</td>
</tr>
<tr>
<td>SP_Rank</td>
<td>0.533</td>
<td>yes</td>
</tr>
</tbody>
</table>
**Research Question 2**: What are the important features to be considered in a search profile?

<table>
<thead>
<tr>
<th>Feature</th>
<th>avg. $\text{NDCG}$</th>
<th>$p$-value $&lt; 0.05$ (vs. B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.5217</td>
<td>—</td>
</tr>
<tr>
<td>Q</td>
<td>0.5259</td>
<td>no</td>
</tr>
<tr>
<td>SP.Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q+T</td>
<td>0.5449</td>
<td>yes</td>
</tr>
<tr>
<td>Q+T+A</td>
<td>0.5453</td>
<td>yes</td>
</tr>
<tr>
<td>Q</td>
<td>0.5155</td>
<td>no</td>
</tr>
<tr>
<td>SP.Rank</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q+T</td>
<td>0.5328</td>
<td>yes</td>
</tr>
<tr>
<td>Q+T+A</td>
<td>0.5334</td>
<td>yes</td>
</tr>
</tbody>
</table>
Research Question 3: *Is there any difference between active and inactive users?*

<table>
<thead>
<tr>
<th></th>
<th>all users</th>
<th>active users</th>
<th>inactive users</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>avg. NDCG</td>
<td>p-value&lt;0.05</td>
<td>avg. NDCG</td>
</tr>
<tr>
<td>B</td>
<td>0.522</td>
<td>—</td>
<td>0.522</td>
</tr>
<tr>
<td>SP_Score</td>
<td>0.545</td>
<td>yes</td>
<td>0.588</td>
</tr>
<tr>
<td>SP_Rank</td>
<td>0.533</td>
<td>yes</td>
<td>0.573</td>
</tr>
</tbody>
</table>

![Graphs showing NDCG distribution for active and inactive users](image-url)
Research Question 4: How many search queries are needed when building a search profile in order to observe quality improvements?
Research Question 5: How much time should the historical information span in order to produce high-quality recommendations? How does the quality vary with the increase in time span?

<table>
<thead>
<tr>
<th></th>
<th>1 month avg. NDCG (vs. B)</th>
<th>2 months avg. NDCG (vs. B)</th>
<th>3 months avg. NDCG (vs. B)</th>
<th>4 months avg. NDCG (vs. B)</th>
<th>5 months avg. NDCG (vs. B)</th>
<th>6 months avg. NDCG (vs. B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value &lt;0.05</td>
<td>p-value &lt;0.05</td>
<td>p-value &lt;0.05</td>
<td>p-value &lt;0.05</td>
<td>p-value &lt;0.05</td>
<td>p-value &lt;0.05</td>
</tr>
<tr>
<td>B</td>
<td>0.522</td>
<td>0.522</td>
<td>0.522</td>
<td>0.522</td>
<td>0.522</td>
<td>0.522</td>
</tr>
<tr>
<td>SP.Score</td>
<td>0.540 yes</td>
<td>0.543 yes</td>
<td>0.545 yes</td>
<td>0.546 yes</td>
<td>0.548 yes</td>
<td>0.549 yes</td>
</tr>
<tr>
<td>SP.Rank</td>
<td>0.524 yes</td>
<td>0.530 yes</td>
<td>0.533 yes</td>
<td>0.534 yes</td>
<td>0.536 yes</td>
<td>0.537 yes</td>
</tr>
</tbody>
</table>

$p$-value <0.05

<table>
<thead>
<tr>
<th></th>
<th>2M vs. 1M</th>
<th>3M vs. 2M</th>
<th>4M vs. 3M</th>
<th>5M vs. 4M</th>
<th>6M vs. 5M</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP.Score</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>SP.Rank</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>
Conclusions

• We addressed the problem of news personalization by leveraging exogenous information extracted from we-search query logs

• We evaluated two strategies of combining news ranking and search ranking

• We provided a thorough experimental evaluation to answer 5 major research questions

• Results overall show that exploiting search profiles leads to considerable improvements
Thanks!

For questions, please refer to the authors of the paper:

Xiao Bai  <xbai@oath.com>
B. Barla Camabzoglu  <barla@berkantbarlacambazoglu.com>
Francesco Gullo  <gullof@acm.org>
Amin Mantrach  <a.mantrach@criteo.com>
Fabrizio Silvestri  <fsilvestri@fb.com>