

# Improving News Personalization through Search Logs\*

Xiao Bai

Yahoo Research, USA

[xbai@oath.com](mailto:xbai@oath.com)

B. Barla Cambazoglu

Independent Researcher, Spain

[barla@berkantbarlacambazoglu.com](mailto:barla@berkantbarlacambazoglu.com)

Francesco Gullo

UniCredit, R&D Dept., Italy

[gullof@acm.org](mailto:gullof@acm.org)

Amin Mantrach

Criteo Research, USA

[a.mantrach@criteo.com](mailto:a.mantrach@criteo.com)

Fabrizio Silvestri

Facebook, UK

[fsilvestri@fb.com](mailto: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

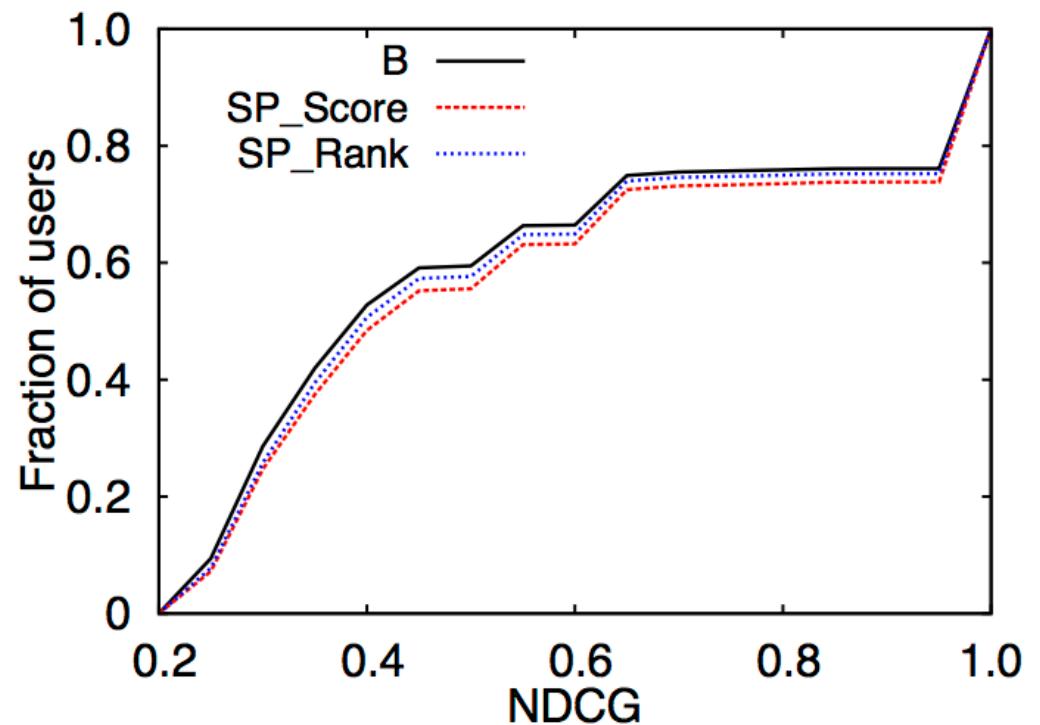
# Experiments - Setting

- 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

# Experiment – Results (1)

**Research Question 1:** *Do search profiles improve the quality of news personalization?*

	avg. $NDCG$	$p\text{-value} < 0.05$ (vs. B)
B	0.522	—
SP_Score	0.545	yes
SP_Rank	0.533	yes



# Experiment – Results (2)

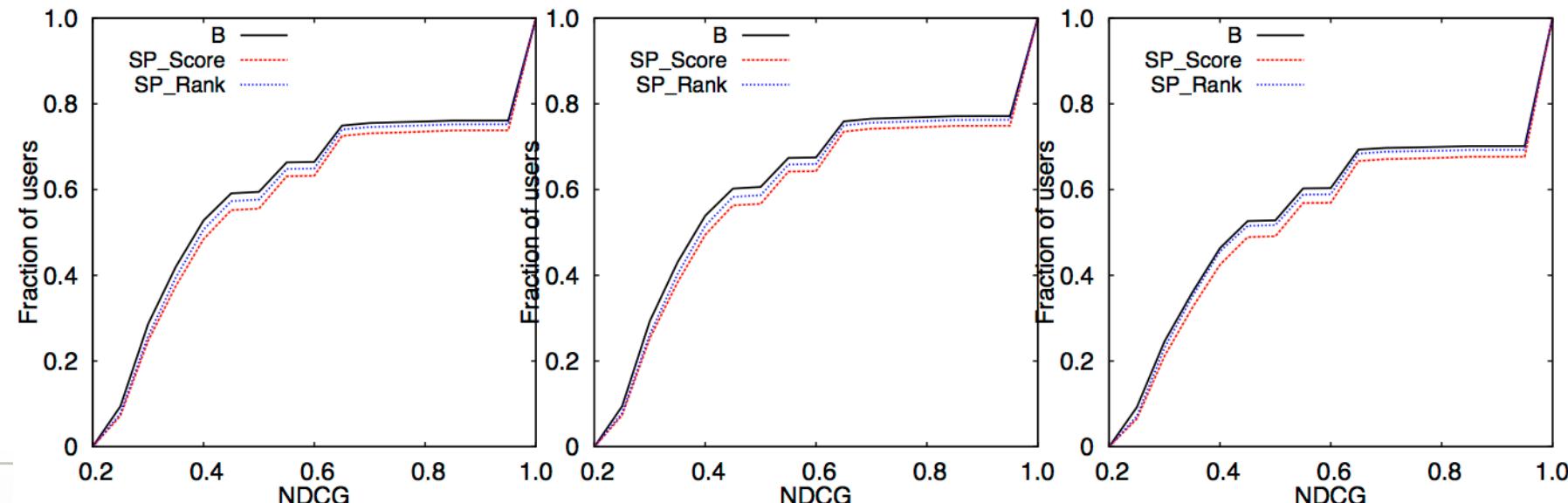
**Research Question 2:** *What are the important features to be considered in a search profile?*

		avg. <i>NDCG</i>	<i>p</i> -value<0.05 (vs. B)
	B	0.5217	—
	Q	0.5259	no
SP_Score	Q+T	0.5449	yes
	Q+T+A	0.5453	yes
	Q	0.5155	no
SP_Rank	Q+T	0.5328	yes
	Q+T+A	0.5334	yes

# Experiment – Results (3)

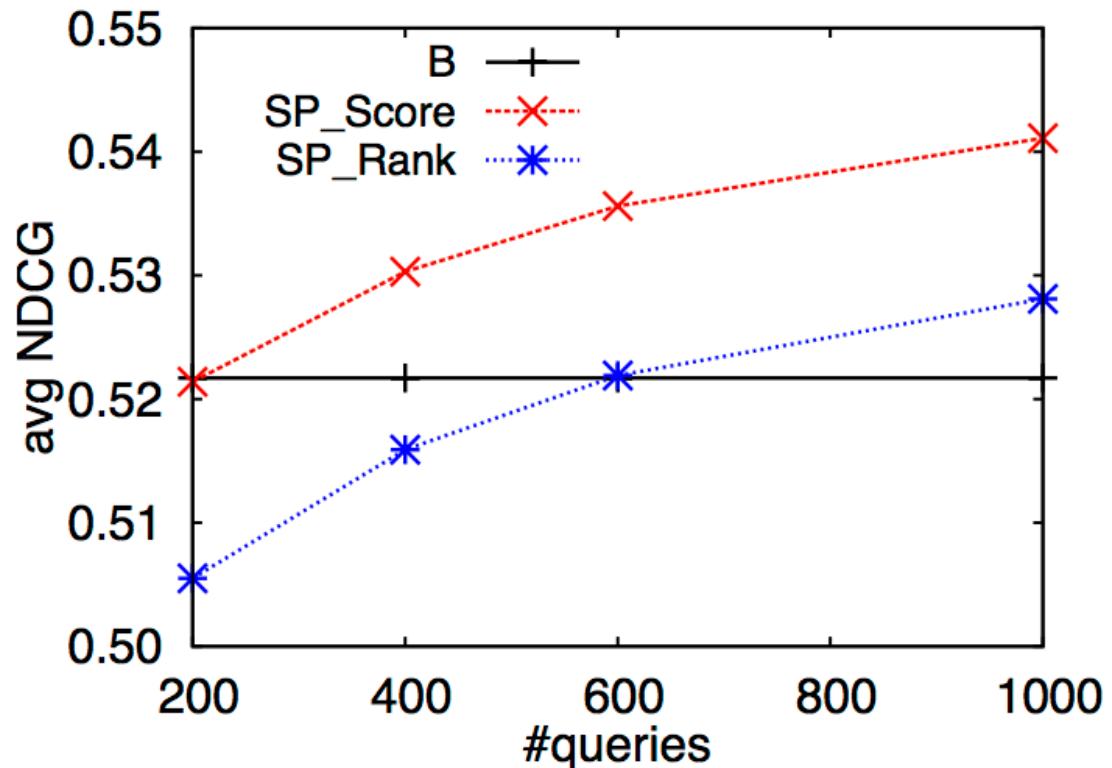
**Research Question 3:** *Is there any difference between active and inactive users?*

	all users		active users		inactive users	
	avg. <i>NDCG</i>	<i>p-value</i> <0.05 (vs. B)	avg. <i>NDCG</i>	<i>p-value</i> <0.05 (vs. B)	avg. <i>NDCG</i>	<i>p-value</i> <0.05 (vs. B)
B	0.522	—	0.522	—	0.522	—
SP_Score	0.545	yes	0.588	yes	0.538	yes
SP_Rank	0.533	yes	0.573	yes	0.526	yes



# Experiment – Results (4)

**Research Question 4:** *How many search queries are needed when building a search profile in order to observe quality improvements?*



# Experiment – Results (5)

**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?*

1 month		2 months		3 months		4 months		5 months		6 months		
	p-value											
	avg.	<0.05		avg.	<0.05		avg.	<0.05		avg.	<0.05	
B	0.522	—	0.522	—	0.522	—	0.522	—	0.522	—	0.522	—
SP_Score	0.540	yes	0.543	yes	0.545	yes	0.546	yes	0.548	yes	0.549	yes
SP_Rank	0.524	yes	0.530	yes	0.533	yes	0.534	yes	0.536	yes	0.537	yes

*p*-value<0.05

	2M vs. 1M	3M vs. 2M	4M vs. 3M	5M vs. 4M	6M vs. 5M
SP_Score	yes	yes	no	no	no
SP_Rank	yes	yes	no	no	no

# 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](mailto:xbai@oath.com)>

B. Barla Camabzoglu <[barla@berkantbarlacambazoglu.com](mailto:barla@berkantbarlacambazoglu.com)>

Francesco Gullo <[gullof@acm.org](mailto:gullof@acm.org)>

Amin Mantrach <[a.mantrach@criteo.com](mailto:a.mantrach@criteo.com)>

Fabrizio Silvestri <[fsilvestri@fb.com](mailto:fsilvestri@fb.com)>