Probabilistic Field Mapping for Product Search

Aman Berhane Ghirmatsion and **Krisztian Balog**
University of Stavanger

CLEF 2015 | Toulouse, France, 2015

Task and Goals

- Product search task of LL4IR
  - Given a keyword query, return a ranked list of products from a product catalog
  - Products are described by a number of attributes (name, brand, categories, characters, etc.)
  - Historical click and query information is available
- Specific goals
  - Establish a probabilistic mapping from query terms to document terms
  - Exploit historical click information

Retrieval Framework

- Requirements
  - Fielded document representation
  - Term-specific weighting in a principled way

Retrieval Framework

Probabilistic Retrieval Model for Semistructured Data (PRMS) [1]

Document language model

$$P(t|\theta_d) = \sum_{f \in F} P(f|t) P(t|\theta_{d_f})$$

Field-mapping probability

Field-specific document language model

[1] J. Kim, X. Xue, and W. B. Croft. A probabilistic retrieval model for semistructured data. In *Proceedings of ECIR 2009*.

Retrieval Framework

Probabilistic Retrieval Model for Semistructured Data (PRMS) [1]

Document language model

$$P(t|\theta_d) = \sum_{f \in F} P(f|t) P(t|\theta_{d_f})$$

Field-mapping probability

Term-field mapping

Field prior

$$P(f|t) = \frac{P(t|f)P(f)}{P(t)} = \frac{P(t|f)P(f)}{\sum_{f' \in F} P(t|f')P(f')}$$

[1] J. Kim, X. Xue, and W. B. Croft. A probabilistic retrieval model for semistructured data. In *Proceedings of ECIR 2009*.

Two Components

- **P(f)** field prior
  - Uniform
  - Set proportional to individual field performance (NDCG), based on historical CTR
- **P(t|f)** probability of a term given a field
  - Uniform
  - Based on term statistics (as in the original PRMS)

Our Methods

- Three specific instantiations of PRMS

	$P(f)$	$P(t\|f)$
Method 1	$\propto \mathrm{NDCG}_f$	$\frac{1}{\|F\|}$
Method 2	$\frac{1}{\|F\|}$	$P_{ML}(t\|C_f)$
Method 3	$\propto \mathrm{NDCG}_f$	$P_{ML}(t\|C_f)$

Fields with est. importance

Field	Description	NDCG
brand	Product's brand	0.0684
category	Product categories	0.4305
characters	Characters associated with the product	0.3792
contents	Catch-all contents field	0.6987
description	Full textual description	0.3919
product_name	Product's name	0.5632
short_description	Short textual description	0.2986
queries	Queries that led to the product	

Results

	Round #1				Round #2			
	Outcome	%W	%L	%T	Outcome	%W	%L	%T
Method 1	0.2827	7.7	19.6	72.7	0.4118	11.5	16.4	72.1
Method 2	**0.3413**	**9.8**	18.9	71.3	0.4389	10.4	13.3	76.2
Method 3	0.3277	8.7	**17.9**	73.4	**0.4795**	**10.7**	**11.6**	77.7

Results

	Round #1				Round #2			
	Outcome	%W	%L	%T	Outcome	%W	%L	%T
Method 1	0.2827	7.7	19.6	72.7	0.4118	11.5	16.4	72.1
Method 2	**0.3413**	**9.8**	18.9	71.3	0.4389	10.4	13.3	76.2
Method 3	0.3277	8.7	**17.9**	73.4	**0.4795**	**10.7**	**11.6**	77.7

In over **70% of the cases** there is a **tie with the production system**. This is the same for all methods.

Results

	Round #1				Round #2			
	Outcome	%W	%L	%T	Outcome	%W	%L	%T
Method 1	0.2827	7.7	19.6	72.7	0.4118	11.5	16.4	72.1
Method 2	**0.3413**	**9.8**	18.9	71.3	0.4389	10.4	13.3	76.2
Method 3	0.3277	8.7	**17.9**	73.4	**0.4795**	**10.7**	**11.6**	77.7

Term-specific field mapping is beneficial.

Method 1 vs. Methods 2 and 3.

Results

	Round #1				Round #2			
	Outcome	%W	%L	%T	Outcome	%W	%L	%T
Method 1	0.2827	7.7	19.6	72.7	0.4118	11.5	16.4	72.1
Method 2	**0.3413**	**9.8**	18.9	71.3	0.4389	10.4	13.3	76.2
Method 3	0.3277	8.7	**17.9**	73.4	**0.4795**	**10.7**	**11.6**	77.7

Mixed results for non-uniform field priors.

Method 2 vs. Method 3.

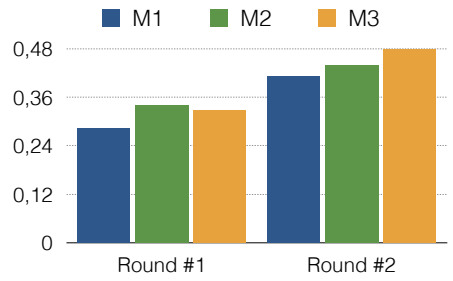
Analysis

How different are the rankings generated by the three methods?

	Method 1	Method 2	Method 3
Method 1		0.867	0.864
Method 2	0.867		0.950
Method 3	0.864	0.950	

Kendall's tau rank correlation.

Stability of Results (?)



Summary

- Best performance among participants :)
  … but still failed to beat the baseline :(
- Lessons learned
  - PRMS offers a principled framework
  - Field-specific term mapping is beneficial
  - Prior field importance is one (promising) way of using historical click information (CTR)
- Future work
  - Better ways to exploit historical queries and clicks
  - Obtain more stable results (run methods for longer periods)