Webis at TREC 2017: Open Search and Core Tracks

Matthias Hagen   Yamen Ajjour   Johannes Kiesel   Payam Adineh   Benno Stein

Bauhaus-Universität Weimar
99421 Weimar, Germany
⟨firstname⟩.⟨lastname⟩@uni-weimar.de

ABSTRACT

We give a brief overview of the Webis group's participation in the TREC 2017 Open Search and Core tracks. Our submission to the Open Search track is similar to our last year's approach, we axiomatically re-rank a BM25-ordered result list to come up with a final ranking. The axiomatic re-ranking idea is also applied in our Core track contribution but with an emphasis on argumentativeness as a not-yet-covered aspect in retrieval.

1 OPEN SEARCH TRACK

The objective of the Open Search track is to rank a small set of given candidate documents (e.g., scientific publications) in return to a given scholarly search query.

Our approach works as follows. We first index the few candidates for a query with Lucene and rank them using Lucene's default BM25 implementation and removing any punctuation from the query. In case that the ranking for the query would be empty, the given candidate documents are just randomly shuffled. The final ranking (i.e., from Lucene or shuffled) is then run through our axiomatic re-ranking pipeline [5] with the axiom weights trained for BM25. We omitted the axioms from our pipeline that are not suited for the scholarly search setup due to missing information (e.g., we omitted the PageRank axiom due to the non-availability of a full citation graph at our site including the candidate documents).

The re-ranking obtained from the combined axioms then forms the submitted ranking similar to our runs for the Web track 2014, the Session tracks 2014–2015 [2, 3], and the Open Search track 2016 [4].

2 CORE TRACK

As for the Core track, our research question was whether including information about the argumentativeness of documents can help in the retrieval process. As the Core track used the New York Times Annotated Corpus and newspaper articles often include argumentation, argumentativeness might be an interesting signal to add to the ranking process in the Core track scenario.

To handle argumentativeness in a retrieval setting, we formulate two basic axioms capturing different aspects of argumentativeness that then should be included in our axiomatic re-ranking pipeline [5]. The first argumentative axiom measures the ratio of argumentative text units among all text units in a document and prefers from a pair of documents the one that has more "arguments." The second axiom measures how close the first occurences of all the query terms in a document are to argumentative text units and prefers from a pair of documents the one where the first occurences are closer to argumentative text units.

Due to some time constraints caused by power outages that crashed some of our experiments, we could not really finish training the weights of our full axiomatic re-ranking approach for the TREC Core track deadline but had to resort to some last-minute submission using a very simplistic pipeline with manually set weights. This is meant to gather some first insights since we obviously could not use the full potential of argumentative axioms. In the simplistic axiomatic re-ranking, we employ the above two argumentativeness axioms and the two more "traditional" axioms TF1 [1] and LB1 [6]. The weights of the four axioms in the re-ranking pipeline were manually set to 0.4 for the argumentativeness axioms and to 0.1 for TF1 and LB1. We plan to include further axioms and a more sophisticated training of the axiom weights in the Core setting but had to leave that open for future research.

REFERENCES

[1] H. Fang, T. Tao, and C. Zhai. A formal study of information retrieval heuristics. In M. Sanderson, K. Järvelin, J. Allan, and P. Bruza, editors, *SIGIR 2004: Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Sheffield, UK, July 25-29, 2004*, pages 49–56. ACM, 2004.

[2] M. Hagen, S. Göring, M. Keil, O. Anifowose, A. Othman, and B. Stein. Webis at TREC 2015: Tasks and Total Recall Tracks. In E. Voorhees and L. Buckland, editors, *24th International Text Retrieval Conference (TREC 15)*, NIST Special Publication. National Institute of Standards and Technology (NIST), Nov. 2015.

[3] M. Hagen, S. Göring, M. Michel, G. Müller, and B. Stein. Webis at TREC 2014: Web, Session, and Contextual Suggestion Tracks. In E. Voorhees and L. Buckland, editors, *23nd International Text Retrieval Conference (TREC 14)*, NIST Special Publication. National Institute of Standards and Technology (NIST), Nov. 2014.

[4] M. Hagen, J. Kiesel, P. Adineh, M. Alahyari, E. Fatehifar, A. Bahrami, P. Fichtl, and B. Stein. Webis at TREC 2016: Tasks, Total Recall, and Open Search Tracks. In E. Voorhees and L. Buckland, editors, *25th International Text Retrieval Conference (TREC 16)*, NIST Special Publication. National Institute of Standards and Technology (NIST), Nov. 2016.

[5] M. Hagen, M. Völske, S. Göring, and B. Stein. Axiomatic Result Re-Ranking. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM 16)*. ACM, Oct. 2016.

[6] Y. Lv and C. Zhai. Lower-bounding term frequency normalization. In C. Macdonald, I. Ounis, and I. Ruthven, editors, *Proceedings of the 20th ACM Conference on Information and Knowledge Management, CIKM 2011, Glasgow, United Kingdom, October 24-28, 2011*, pages 7–16. ACM, 2011.