Citation Analysis for Academic Expert Finding

ABSTRACT

Bla bla bla

Keywords

citation analysis, expert finding

1. INTRODUCTION

What is expert search? Different types of evidence? Why is expert search important? Say that it has received a lot of attention in TREC. Mostly focused on the content-based aspect. Argue for potential of citations. Some work done on including citations, but nothing serious. We examine the use of citations for expert finding on a large scale.

Existing collections do not have citations, so we need a new collection that contains citations. We use the DBLP data set as our document collection, because a subset of it contains citations. To get topics, we use a novel method: taking workshop website descriptions and using the organizers and program committee members as relevance judgments.

Contributions: (1) studying citation analysis for expert finding on large scale, (2) novel way of producing topic set for test collection, (3) comparing different types of citation analysis metrics.

The remainder of this paper is organized as follows. ...

2. RELATED WORK

3. METHODOLOGY

Bla bla bla

3.1 Test collection creation

DBLP data set has is big, has metadata, and has citation information (with some caveats, list them) / Can be used to support expert finding / Need to create topics that can be used for this with relevance assessments in an easy way / Workshops and tutorials have descriptions and associated organizers and PC members / Are used for topic creation, relatively easy process / Relevance assessments can be graded or binary

3.2 Citation analysis metrics

Document-centric metrics / Person-centric metrics / Venue-centric metrics

3.3 Experimental setup

50 workshop and 50 tutorial topics from IR conferences of recent years (SIGIR, CIKM, IIIX, RecSys, DL, ECIR) / Take 25 from each as our training set to optimize parameters / Take remaining 25+25 topics as our test set / We report our metrics on this / MRR for top expert / NDCG@10 for top 10 / Both graded and binary relevance judgments?

4. RESULTS

4.1 Document-centric metrics

4.2 Person-centric metrics

4.3 Venue-centric metrics

5. DISCUSSION & CONCLUSIONS

Bla bla bla

5.1 Future work

More metrics / Manually checking the relevance assessments / Exploring the value of different topic fields / Collecting PDFs of documents

6. REFERENCES