The Impact of Author Ranking in a Library Catalogue

Jaap Kamps
University of Amsterdam, The Netherlands

ABSTRACT

The field of information retrieval has witnessed over 50 years of research on retrieval methods for metadata descriptions and controlled indexing languages, the prototypical example being the library catalogue. It seems only natural to resort to additional data for improving book retrieval, such as the text of the book in whole or in part (table of contents, abstract) or contributed social data acquired through crowdsourcing social cataloguing sites like LibraryThing. Without denying the potential value of such additional data, we want to challenge the underlying assumption that applying novel retrieval methods to traditional book descriptions cannot improve book retrieval. Specifically, this paper investigates the effectiveness of author rankings in a library catalogue. We show that a standard retrieval model results in a book ranking that meets and exceeds the effectiveness of catalogue systems. We show that using expert finding methods we also can obtain effective author rankings that complement the traditional book rankings. Moreover, ranking books on author scores leads to substantial and significant improvements over the original book rankings. If we base our book ranking on the combination of the author scores and the book scores we see no further improvements. Hence our results clearly demonstrate the importance of author ranking for retrieving library catalogue records: authors capture an important aspect of relevance and one that is not obvious to those unfamiliar with specific area of interest.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Search process, Selection process*; H.3.7 [**Information Storage and Retrieval**]: Digital Libraries—*Systems Issues, User Issues*

General Terms

Experimentation, Measurement, Theory

Keywords

Library catalogue, Expert finding, Entity ranking

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
BooksOnline'11, October 24, 2011, Glasgow, Scotland, UK.
Copyright 2011 ACM 978-1-4503-0961-5/11/10 ...$10.00.

1. INTRODUCTION

The field of information retrieval (IR) is rooted on the field of library science, where the problem of bringing searchers and information together has been studied for centuries. Yet at the same time few, if any, of the advances of modern IR have been applied within library systems. Searching in a library catalogue, as students and scholars around the globe experience on a daily basis, can be a disappointing experience due to the relatively basic search systems working on short bibliographic descriptions. This paper investigates whether some recent IR methods, in particular expert or entity ranking [1, 7, 8], can be fruitfully applied to a library catalogue.

An encyclopedic overview of online public access catalogs (OPACs), and their evolution over time, is beyond the scope of this paper. Borgman [4, 5] gives an excellent overview of the development of the OPAC and its remaining problems up to the mid 90's. Since the rise of the Web and its gatekeepers, the Internet search engines, traditional OPACs have rapidly lost ground in particular for subject access [10]. Unprecedented requests for radical changes are also voiced in the context of the next generation of cataloguing rules [6]. It remains an open question whether the traditional paradigm of the OPAC, based on control as a central principle, can successfully be combined with the emerging paradigm of Web Search, based on uncertainty and statistics. A recent example of an innovative library catalog is the parametric or faceted search as pioneered at the North Carolina State University NCSU [11]. The current paper has a more limited scope, and focuses on the ability of modern expert or entity ranking methods to generate an author ranking for subject search based on library catalogue descriptions, and on the usefulness of such an author ranking relative to the traditional book ranking.

Our research is motivated by a concrete use case. A substantial fraction of users of library catalogues has only limited understanding of the topic of request—the very reason they search for information on this topic [3]. Such a searcher, think of a student new to a field, often has substantial difficulty in selecting the 'right' publications from a ranked list. Knowledgeable searchers, think of a scholar searching for literature in their field of expertise, often immediately see what's good and what's not so good. One reason is that expert searchers can exploit additional retrieval cues that are not meaningful to uninitiated searchers. One clear example is recognizing important authors on the topic at hand, and we will try to exploit this impact of the author in this paper. We expect that the importance of such additional

```
<kbs xmlns:xsi="http://www.w3.org/1999/XMLSchema-instance">
  <record>
    <identifier>http://resolver.kb.nl/resolve?urn=PPN:190806788</identifier>
    <title>Advances in information retrieval : recent research from the Cen-
ter for Intelligent Information Retrieval / ed. by W. Bruce Croft</title>
    <contributor>Croft, W. Bruce</contributor>
    <contributor>Center for Intelligent Information Retrieval</contributor>
    <publisher>Boston ; Dordrecht [etc.] : Kluwer Academic Publishers</publisher>
    <isPartOf>The Kluwer international series on information retrieval, ISSN 1387-
5264 ; 7</isPartOf>
    <subject type="corpbody">Center for Intelligent Information Re-
trieval</subject>
    <subject type="GOO">Information storage and $retrieval</subject>
    <subject type="GOO">Databanken</subject>
    <subject type="Brinkman">information retrieval</subject>
    <subject type="BCL">06.64</subject>
    <subject type="BCL">54.64</subject>
  </record>
</kbs>
```


Figure 1: Sample book record.

cues is particularly important when there are relatively few textual cues—such as searching short bibliographic descriptions.

Our research questions are the following:

- How does a state-of-the-art retrieval engine compare to a state-of-the-art library catalogue system?
- Using only the information already in the library catalogue, can we produce author rankings? Do these resonate closely with the "relevance" of authors?
- Can book rankings be improved using the author evidence? And if so, how important is the author evidence relative to the original book scores?

The rest of this paper is structured as follows. Next, in Section 2, we discuss the experimental setup: the used data, systems and evaluation. Section 3 discusses our results: the effectiveness of catalogue retrieval, of novel methods for author ranking, and of book retrieval methods based on the author scores. We end in Section 4 by discussing our results and drawing initial conclusions.

2. EXPERIMENTAL SETUP

In this section, we detail out experimental setup by discussing the used data, the used systems, and the evaluation.

2.1 Data: Library Catalogue

We use the complete catalogue of the *Koninklijke Bibliotheek*, the National Library of the Netherlands. These records are kept in proprietary systems, but have been exported into XML. The resulting set consists of 2,102,357 book records. Figure 1 shows an example record simplified to the fields containing named and subject access points.

2.2 System: Book and Author Ranking

Our main interest is in exploring the effectiveness of derived author rankings, and of book rankings based on the author rankings. Hence we use a standard language model without further tuning to produce an initial book ranking, and focus on deriving an author ranking based on these scores.

We imported all XML records directly into the PF/Tijah system [9]. This allows us to search for records using various IR models in combination with powerful XQuery/NEXI operators in the following ways.

Book Ranking For our own book rankings we use the multinomial language model with linear smoothing, with default settings.

Specifically, for a collection D, document d and query q:

$$P(d|q) = P(d) \cdot \prod_{t \in q} \left((1-\lambda) \cdot P(t|D) + \lambda \cdot P(t|d) \right),$$

where

$$\begin{aligned} P(t|d) &= \frac{tf_{t,d}}{|d|}, \\ P(t|D) &= \frac{\mathsf{doc\_freq}(t,D)}{\sum_{t' \in D} \mathsf{doc\_freq}(t',D)}, \\ P(d) &= \frac{|d|}{\sum_{d' \in D} |d'|}. \end{aligned}$$

The standard value for the smoothing parameter λ is 0.2. We retrieve maximally 150 books.

Author Ranking Based on the ranked books, we explore different author rankings based on the author frequency in the top 150 results, and the highest, average, and sum of book scores. That is, we derive here the author ranking solely from the book ranking above by propagating the estimated relevance of retrieved books to their respective authors.

Specifically, for an author a, document d and query q:

$$\begin{aligned} P_{\mathsf{author\_max}}(a|q) &= \max_{d: a \in \mathsf{author}(d)} P(d|q), \\ P_{\mathsf{author\_mean}}(a|q) &= \frac{\sum_{d: a \in \mathsf{author}(d)} P(d|q)}{|\{d : a \in \mathsf{author}(d)\}|}, \\ P_{\mathsf{author\_sum}}(a|q) &= \sum_{d: a \in \mathsf{author}(d)} P(d|q), \end{aligned}$$

with $\mathsf{author}(d)$ denoting the author(s) of document d.

Ranking Books on Author Scores Having obtained various author rankings, we investigate what happens if we use the author scores to rank books, based on the sum of the author scores. That is, we derive here a new book ranking solely from the author ranking above by propagating the author's relevance scores to the books they have written. The original book ranking is used only indirectly, since it is the basis for the author ranking above.

Specifically, for a document d, author a, and query q:

$$\begin{aligned} P_{\mathsf{book\_max}}(d|q) &= \sum_{a \in \mathsf{author}(d)} P_{\mathsf{author\_max}}(a|q), \\ P_{\mathsf{book\_mean}}(d|q) &= \sum_{a \in \mathsf{author}(d)} P_{\mathsf{author\_mean}}(a|q), \\ P_{\mathsf{book\_sum}}(d|q) &= \sum_{a \in \mathsf{author}(d)} P_{\mathsf{author\_sum}}(a|q). \end{aligned}$$

Ranking Books on Mixtures Finally, we explore mixtures of the author and book scores. That is, here we use both the original book ranking and the derived author ranking.

Table 1: Topics and statistics.

Query	Books Jud.	Books Rel.	Authors Jud.	Authors Rel.
1 amerika geschiedenis	119	15	67	4
2 amerika politiek	110	6	69	1
3 auteursrecht software	101	13	68	4
4 digitale bibliotheek	85	0	–	–
5 economie crisis	143	17	106	13
6 filosofie plato	140	37	79	7
7 geschiedenis rome	134	18	83	7
8 information retrieval library	128	12	87	4
9 information retrieval	125	25	86	9
10 tweede wereldoorlog	125	41	59	7

Specifically, for a document d, author a, and query q:

$$\begin{aligned} P_{\mathsf{mix\_max}}(d|q) &= \mu \cdot P(d|q) + (1-\mu) \cdot P_{\mathsf{book\_max}}(d|q), \\ P_{\mathsf{mix\_mean}}(d|q) &= \mu \cdot P(d|q) + (1-\mu) \cdot P_{\mathsf{book\_mean}}(d|q), \\ P_{\mathsf{mix\_sum}}(d|q) &= \mu \cdot P(d|q) + (1-\mu) \cdot P_{\mathsf{book\_sum}}(d|q), \end{aligned}$$

where μ is a linear interpolation factor in the range $[0, 1]$. With $\mu = 1$ we obtain the original book ranking, and with $\mu = 0$ we obtain the book rankings based on the author scores.

We also compare our rankings to the on-line public access catalogue (OPAC) of the Dutch National Library, available at `http://opc4.kb.nl/`, which ranks by default on *relevance*, or alternatively on *recency*—generally not very effective for informational search.

2.3 Evaluation: Informational Search

We build a small test set of 10 ad hoc search topics, reflecting comprehensive search tasks typical for university students, with a short search request of 2-3 words. Table 1 lists the 10 topics and some statistics about the judged and relevant books and derived judged and relevant authors.

We made shallow pools of the KB catalogue run (ordered on relevance) and our own index (language model run). The pools consist of up to 150 books, and are judged by a single assessor. As it turned out, for one of the topics (topic 4, 'digital library' in Dutch) the search request was too general and no relevant books were found. Hence we evaluate over the remaining nine topics, with on average 20 relevant book records. We also look at the relevance of authors and derive author-judgments by defining a relevant author as the author of a relevant book.

3. EXPERIMENTAL RESULTS

In this section, we discuss the result of our experiments on the effectiveness of catalogue retrieval, of novel methods for author ranking, and of book retrieval methods based on the author scores.

3.1 Catalogue Search

How effective is our language model in comparison to the library's public access catalogue? Table 2 shows that the recency ranking is not very effective for informational search. The relevance ranking fares much better—the information request were relatively specific and the and-ish ranking of the OPAC obtains a good precision. The language model scores somewhat better, even though standard settings were used. Figure 2 shows the precision-recall curves of the book rankings.

Table 2: Book record rankings (best scores in bold).

Run	map	P@1	P@5	P@10
Catalogue recency	0.0948	0.2697	0.1111	0.1222
Catalogue relevance	0.2484	0.6177	0.3778	**0.3444**
Language Model	**0.2536**	**0.7432**	**0.4222**	0.3000

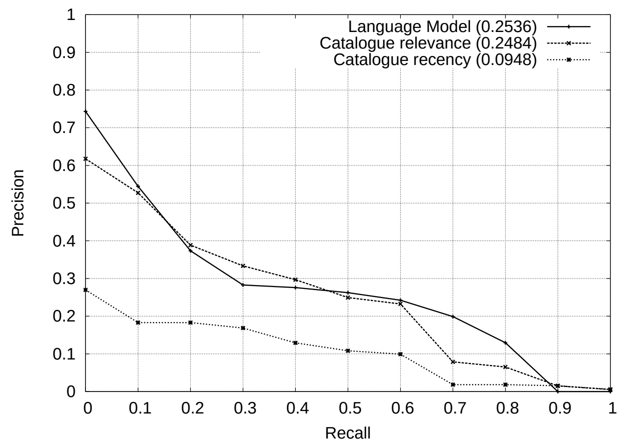

Figure 2: Book rankings: Interpolated precision over recall.

Table 3: Author rankings based on book record scores (best scores in bold).

Run	map	P@1	P@5	P@10
Author frequency	0.2406	0.5472	0.2444	0.2111
Mean book score	0.4124	0.8333	0.3556	0.2778
Highest book score	0.4206	**0.8407**	**0.4222**	0.2778
Sum of book scores	**0.4459**	0.8258	0.3778	**0.3000**

3.2 Author Ranking

Can we use the book ranking as a basis to rank the authors? Table 3 shows the resulting author rankings. As a baseline we use the ranking of authors on their frequency in the top 150—similar to the faceted (refined or parametric) search options that some advanced catalogue systems experiment with. The author rankings based on book scores are significantly better ($p < 0.01$, one-tailed, bootstrap test) than the frequency based author ranking. The ranking using the highest book gets the highest scores at rank 1 and 5, but the ranking using the sum of book scores gets the highest overall scores. Figure 3 shows the precision-recall curves for the author rankings.

3.3 Book Ranking based on Author Ranking

What happens if we now rank the book records on the author scores? Table 4 shows the book rankings based on the sum of author scores. We restrict our attention to the sum of author scores because experiments with the mean or highest author score, not reported here, showed inferior results. As a baseline we include the best system based on

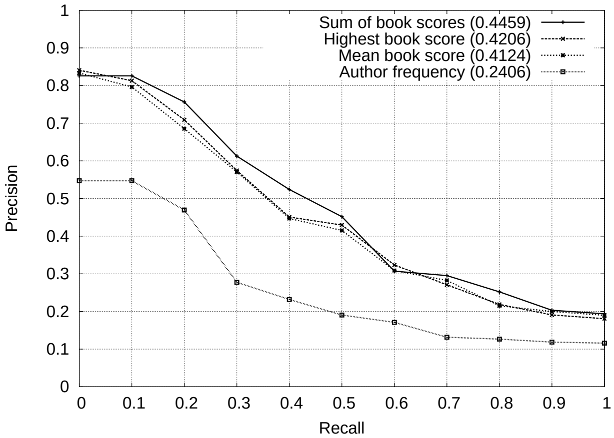

Figure 3: Author rankings: Interpolated precision over recall.

Table 4: Book record rankings based on author scores (best scores in bold).

Run	map	P@1	P@5	P@10
Language Model	0.2536	0.7432	**0.4222**	0.3000
Author frequency	0.2956	0.5347	0.3778	0.3889
Mean book score	0.2787	**0.8259**	0.4000	0.2889
Highest book score	0.2830	0.8120	**0.4222**	0.3333
Sum of book scores	**0.3393**	0.7544	**0.4222**	**0.4000**

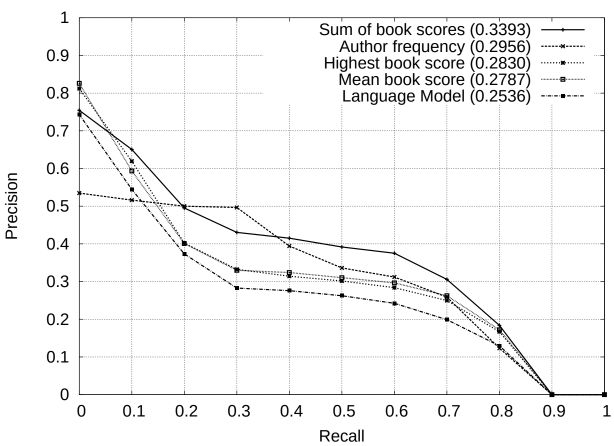

Figure 4: Book rankings based on author scores: Interpolated precision over recall.

the catalogue data, our language model run. We see that this leads to an increase in overall precision (map). The improvement for the sum variant is statistically significant ($p < 0.05$, one-tailed, bootstrap test). This results indicates that the authors indeed convey important information. Figure 4 shows the precision-recall curves for the book rankings based solely on the author scores, relative to the performance of our language model on the book records.

3.4 Combining Book and Author Scores

Can we fruitfully combine author and book scores? Figure 5 plots the effect of combining the author-based book scores

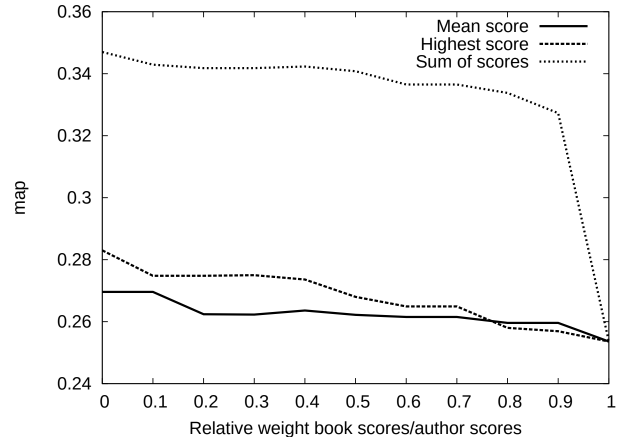

Figure 5: Combination of book and author scores.

with the original book scores. What we see is that all three variants peak at weight zero: the book ranking is solely based on the author ranking and there is no additional benefit of including the original book scores. As a result, the optimal 'combination' score is identical to the book ranking based on the author scores as shown above in Table 4.

4. CONCLUSIONS

In this paper we investigated the effectiveness of author rankings in a library catalogue, aiming to improve book retrieval using no more than the information already available in the library catalogue records. Our main findings are the following. First, standard IR models provide effective document ranking that meets and exceeds the effectiveness of catalogue systems. Second, using expert finding methods we also can obtain effective author rankings, which are substantially and significantly better than simple frequency counting. Third, ranking documents on author scores leads to substantial further improvements in the document rankings, the best scoring one being a significant improvement over the strong baseline of the IR model. Fourth, combining the author scores with the original document scores does not lead to further improvement: the best scoring document ranking is solely based on the author ranking. This signals that authors capture indeed an important aspect of relevance—and one that is not obvious to those unfamiliar with specific area of interest.

In the above experiments we used an author ranking based on a straightforward document model for authors, more sophisticated models as proposed for expert finding and entity retrieval may lead to further improvements [2, 12]. Apart from improving the document ranking in a text scarce environment, the author rankings themselves are a welcome addition to traditional catalogues providing insight into the result set. Having established to importance of author information directly suggests exploring further aspects of the structured library descriptions, such as the classification, subject heading, publisher, or year of publication, in similar ways.

Our research was motivated by a concrete use case of uninitiated searchers, such as students new to a field, consulting the literature being unaware of the reputations of its authors. We have demonstrated by system evaluation that we can capture this particular aspect, the author's reputation as an expert for the subject at hand, using expert or entity search methods, and that this is useful for both the tasks of author ranking and of book ranking. In future research should we hope to investigate how the proposed author ranking impacts the search experience in a user study or operational experiment. In addition, with a more substantial evaluation benchmark, further refinement of the exact ranking model and its parameters can also yield further improvements.

Acknowledgments

Thanks to the National Library of the Netherlands for making their entire catalogue available for research purposes, and thanks to Lysander Vogelzang for running the experiments on the author ranking as part of his Master's thesis [13].

This work was generously supported by the Netherlands Organization for Scientific Research (NWO, grants # 612.-066.513, 639.072.601, and 640.005.001) and by the European Community's Seventh Framework Programme (FP7 2007/2013, Grant Agreement 270404).

REFERENCES

[1] P. Bailey, N. Craswell, I. Soboroff, and A. P. de Vries. The CSIRO enterprise search test collection. *SIGIR Forum*, 41(2):42–45, 2007.

[2] K. Balog, L. Azzopardi, and M. de Rijke. A language modeling framework for expert finding. *Information Processing and Management*, 45:1–19, 2009.

[3] N. J. Belkin. Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, 5:133–143, 1980.

[4] C. L. Borgman. Why are online catalogs hard to use? lessons learned from information-retrieval studies. *JASIS*, 37:387–400, 1986.

[5] C. L. Borgman. Why are online catalogs still hard to use? *JASIS*, 47:493–503, 1996.

[6] K. Coyle and D. Hillmann. Resource description and access (RDA): Cataloging rules for the 20th century. *D-Lib Magazine*, 1/2, 2007.

[7] N. Craswell, A. P. de Vries, and I. Soboroff. Overview of the TREC-2005 enterprise track. In E. M. Voorhees and L. P. Buckland, editors, *The Fourteenth Text REtrieval Conference (TREC 2005)*. National Institute of Standards and Technology. NIST Special Publication, 2006.

[8] A. P. de Vries, A.-M. Vercoustre, J. A. Thom, N. Craswell, and M. Lalmas. Overview of the INEX 2007 entity ranking track. In N. Fuhr, M. Lalmas, A. Trotman, and J. Kamps, editors, *Focused access to XML documents (INEX 2007)*, volume 4862 of *Lecture Notes in Computer Science*, pages 245–251. Springer, 2008.

[9] D. Hiemstra, H. Rode, R. van Os, and J. Flokstra. PF/Tijah: text search in an XML database system. In *Proceedings of the 2nd International Workshop on Open Source Information Retrieval (OSIR)*, pages 12–17, 2006.

[10] K. Markey. The online library catalog: Paradise lost and paradise regained? *D-Lib Magazine*, 13(1/2), 2007.

[11] NCSU. Libraries online catalog, 2011. `http://www.lib.ncsu.edu/catalog/`.

[12] P. Serdyukov and D. Hiemstra. Modeling documents as mixtures of persons for expert finding. In C. Macdonald, I. Ounis, V. Plachouras, I. Rutven, and R. W. White, editors, *Advances in Information Retrieval: 30th European Conference on IR Research (ECIR 2008)*, volume 4956 of *LNCS*, pages 309–320. Springer, 2008.

[13] L. Vogelzang. Improving access to digital libaries by using metadata. Master's thesis, Information Science, University of Amsterdam, 2009. `http://dare.uva.nl/en/scriptie/297597`.