Sponsored vs. Organic (Research Paper) Recommendations and the Impact of Labeling

Joeran Beel, Stefan Langer, and Marcel Genzmehr

Docear, Germany
{beel,langer,genzmehr}@docear.org

Abstract. In this paper we show that organic recommendations are preferred over commercial recommendations even when they point to the same freely downloadable research papers. Simply the fact that users perceive recommendations as commercial decreased their willingness to accept them. It is further shown that the exact labeling of recommendations matters. For instance, recommendations labeled as 'advertisement' performed worse than those labeled as 'sponsored'. Similarly, recommendations labeled as '*Free* Research Papers' performed better than those labeled as 'Research Papers'. However, whatever the differences between the labels were – the best performing recommendations were those with no label at all.

**Keywords:** recommender systems, organic search, sponsored search, labeling.

1 Introduction

In the Web community there is lots of discussion about organic and sponsored search. 'Organic search' is the classic search where users enter search terms and search engines return a list of relevant web pages. 'Sponsored search' describes additional 'results' that are often shown beside the organic results. Usually these results are related to the search terms but companies pay for them to be displayed (in other words, 'sponsored search' is a nice paraphrase for personalized advertisement). While typical online advertisement has click-through rates (CTR) around 0.5% [1], sponsored search achieves CTRs around 2% and sometimes even more than 30% [2]. CTR is a common performance measure in online advertisement. It describes how many ads were clicked relative to the delivered ones. For instance, if 1,000 ads were delivered, and users clicked 61 of them, CTR was 6.1%. The higher the CTR the better is the algorithm behind the search results.

In academia, there are several academic recommender systems which typically only show organic recommendations [3, 4]. However, we were interested which CTR was to expect for sponsored recommendations in academia and more importantly, how much, or how little, users would like recommendations in general that were displayed for profit-making.

2   Methodology

Our academic literature management software 'Docear' [6] features a research paper recommender system [5]. Every third start Docear displays ten recommendations that can be freely downloaded (Figure 1). We modified Docear's recommender system and analyzed the effects of the modifications on click-through rates (overall, 22,452 recommendations were delivered to 587 users). Modifications were related to a label describing the nature of the recommendations (organic or commercial) and the way of presenting recommendations (Figure 1). More information on the recommender system can be found in [5, 6].
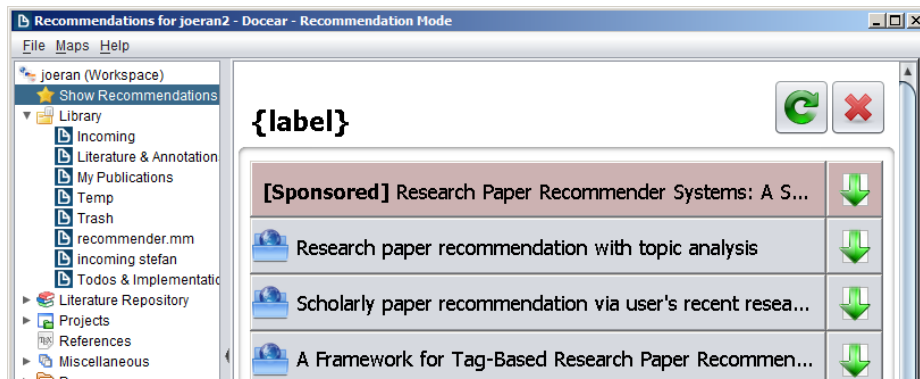
Fig. 1. Recommendations in Docear (top recommendation 'sponsored' and highlighted)

Recommendations in Docear were 'labeled' to explain the 'nature' of the recommendations (Figure 1). The 'basic' label was 'Research Papers'. We modified this label for each user by randomly choosing whether to add a prefix such as 'Free' or 'Free Full-text' (Table 1) or a suffix such as '(Advertisement)' or '(Sponsored)' which resulted in labels like 'Free Research Papers', 'Research Papers from our partners', or 'Free Full-text Research Papers (Sponsored)'. When a suffix was chosen, user must have assumed that the recommendations had a commercial background. When no suffix was chosen, users must have assumed that recommendations were organic. In addition, when no suffix was chosen it was randomly chosen whether to mark the first recommendation as '[Sponsored]' and whether to highlight this recommendation or not (Figure 1). Whatever label was displayed, recommendations were always calculated with the same algorithms and always linked to freely downloadable PDFs.

Table 1. Labels for the recommendations

Prefix				Suffix		
Free	Free Full-text	Full-text	None	(Sponsored)	(Advertisement)	From our partners

We selected two metrics to measure the effectiveness of recommendations and determine differences between the labels. With click-through rate (CTR) we measured how many recommendations out of the displayed ones were clicked overall. For instance, if 1,000 recommendations with a certain label were shown and 50 were clicked, CTR was 5%. If CTR for recommendations with another label was, for instance, 3.2%, the first label performed better. CTR is a common measure on advertisement but it suffers from one problem, especially when recommendations of only a few users are analyzed. In this case, a few users could spoil the results. For instance, one user receiving and clicking many recommendations would strongly increase overall CTR, although maybe all other users hardly clicked on any recommendations. Therefore, we also used mean average precision (MAP) for users' click-through rates. That means, for each user we calculated his average CTR and then we calculated the mean CTR over all users. For instance, if one user had seen 50 recommendations and clicked all of them, and 95 other users had each seen 10 recommendations but clicked none, CTR for the first user was 100% but CTR for the 95 others were 0% each. Hence, MAP was $\frac{100\%+0\%+0\%+\cdots+0\%}{96} = 1.04\%$.

3   Results

Based on CTR organic recommendations clearly outperform commercial ones with a CTR of 8.86% vs. 5.86% (Figure 2, blue line). This is probably what most people would expect. However, it is still interesting to have it quantified that only because recommendations are labeled as some kind of commercial, users are far less likely to click on them. Based on CTR, recommendations with the first recommendation being labeled as '[Sponsored]', but not highlighted, also clearly outperform those being highlighted (8.38% vs. 5.16%). However, the evaluation based on MAP shows a different picture (Figure 2, beige line). Here, organic (MAP=5.21%) and commercial recommendations (4.91%) perform very much alike. In addition, recommendations with the first one being labeled as sponsored *and* being highlighted (MAP=7.47%) outperform those being not highlighted (5.25%). What is evident with both metrics is that completely unlabeled recommendations performed better than all other label variations (CTR=9.87%; MAP=8.76%).
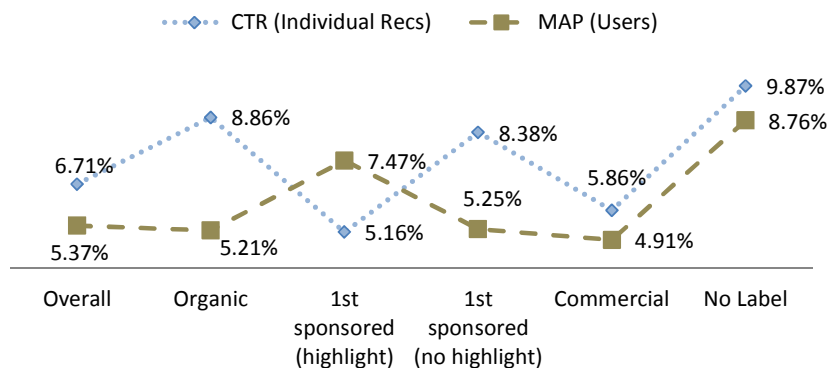

Fig. 2. CTR and MAP of different labels

For organic recommendations the 'free' and 'free full-text' labels clearly outperformed those labels not indicating that the recommended papers were free to download (Figure 3). This is true for both metrics CTR and MAP[1]. However, for commercial recommendations results differed. Here, using no suffix at all (MAP=6.51%; CTR=7.26%) performed better than any of the suffixes. We cannot explain this difference. For suffixes, both CTR and MAP indicate that 'Advertisement' leads to the lowest performance (Figure 4). Based on MAP 'Sponsored' recommendations (5.95%) performed better than 'partner' recommendations (4.85%). Based on CTR, 'partner' recommendations performed better (6.79%) than 'sponsored' ones (5.93%).

Summarized, the most surprising result was that recommendations with no label at all performed best, and that based on MAP commercial and organic recommendations performed about alike. Our study also showed that click-rates on recommendations varied strongly based on how they were labeled (although they were all based on the same algorithms). In particular recommendations labeled as 'advertisement' were least liked by the users. Results based on CTR often contradicted those based on MAP and also using certain prefixes had different effects on commercial and organic recommendations. More research is needed to clarify these contradictions. In some cases a small sample size might have caused the contradictions. For instance, for some labels (e.g. 'Free Research Papers') results were only based on twelve users. However, other results were based on larger samples and still contradict each other.

Open Data. Due to space restrictions, some data and graphs were omitted in this paper. For those readers being interested in more details, e.g. exact numbers of users per label, or validating our research, we publish additional data on http://labs.docear.org.
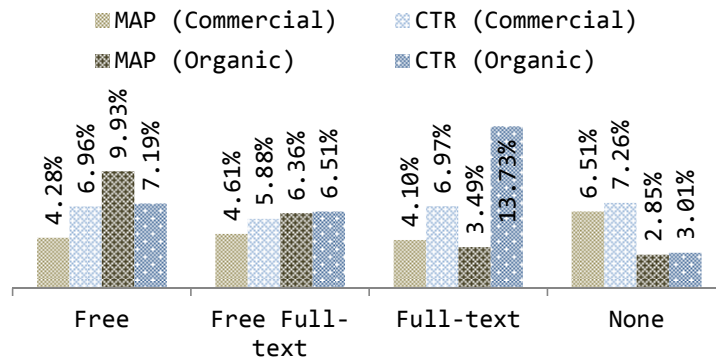
Fig. 3. MAP and CTR for prefixes (commercial and organic)

---

[1] For 'full-text' CTR is an outlier. We investigated the result and found that in this case few users had extremely high CTRs based on few received recommendations they almost all clicked.
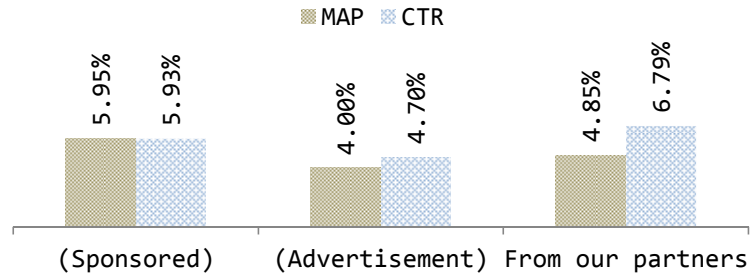
MAP   CTR

5.95% 5.93%   4.00% 4.70%   4.85% 6.79%

(Sponsored)   (Advertisement) From our partners

Fig. 4. MAP and CTR for suffixes (commercial only)

References

1. Manchanda, P., Dubé, J.P., Goh, K.Y., Chintagunta, P.K.: The effect of banner advertising on internet purchasing. Journal of Marketing Research 43 (2006)
2. Schwartz, B.: Google AdWords Click Through Rates: 2% is Average But Double Digits is Great. Search Engine Round Table Blog (2010),
   http://www.seroundtable.com/archives/021514.html
3. Gori, M., Pucci, A.: Research paper recommender systems: A random-walk based approach. In: Proceedings of the International Conference on Web Intelligence (2006)
4. Zhang, M., Wang, W., Li, X.: A Paper Recommender for Scientific Literatures Based on Semantic Concept Similarity. In: Buchanan, G., Masoodian, M., Cunningham, S.J. (eds.) ICADL 2008. LNCS, vol. 5362, pp. 359–362. Springer, Heidelberg (2008)
5. Beel, J., Langer, S., Genzmehr, M., Nürnberger, A.: Introducing Docear's Research Paper Recommender System. In: Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, JCDL (2013)
6. Beel, J., Gipp, B., Langer, S., Genzmehr, M.: Docear: An Academic Literature Suite for Searching, Organizing and Creating Academic Literature. In: Proceedings of the 11th ACM/IEEE Joint Conference on Digital Libraries, pp. 465–466. ACM (2011)