Leading People to Longer Queries in Site Search[*]

Exploring the Query Halo Effect

Djoerd Hiemstra
University of Twente
Enschede, The Netherlands
hiemstra@cs.utwente.nl

Claudia Hauff
Delft University of Technology
Delft, The Netherlands
c.hauff@tudelft.nl

Leif Azzopardi
University of Strathclyde
Glasgow, Scotland
leif.azzopardi@strath.ac.uk

1 INTRODUCTION

People tend to type short queries, however, the belief is that longer queries are more effective. In this paper, we explore the query halo effect, proposed by Agapie et al. [1], where the search box glows and changes colour as the user types. Agapie et al. [1] showed that this halo effect led people to enter longer queries in the context of complex Web search tasks in a laboratory setting. To this end, we consider the following: *1)* Does the query halo effect lead people to longer queries in a natural setting (site search)? and *2)* Does the assumption of longer queries being more effective hold in this setting? We empirically investigated these questions in a 50-day long A/B test setup implemented in a university site search engine.

2 REPRODUCING THE HALO EFFECT

Experiments were run on the site search engine of the University of Twente[1], a federated search engine searching 35 resources including Google's site search, local courses, local news, the telephone directory, the university timetables, as well as results from the university's social media feeds, such as Facebook, Twitter and Flickr. Given a query, the search engine returns ranked resource *blocks* with each block containing up to four (up to seven in the case of images) ranked items; each resource can only contribute a single block to a ranking. Our query log records contain for each query the URLs of the search results that were clicked as well as the block rank of the clicked result. As the vast majority of queries yielded a single click, we computed the block rank MRR (the mean reciprocal rank over all submitted queries in the respective condition) as our system's measure of retrieval performance. The implementation uses the open source federated search engine Searsia[2]. Experiments were run as A/B tests where users were assigned randomly to either the control condition (the standard search box, labelled henceforth as no-halo condition) or the experimental condition (the search box with the query halo effect, labelled as halo condition).

3 RESULTS

Table 1 (top) contains the results of the A/B test. Users in the control condition submit on average 2.18 query terms; users in the halo

condition match this almost perfectly with an average query length of 2.16. Splitting the users according to their experience with the search system – users submitted a single query across the three weeks (*single query users*) vs. users submitting two or more queries (*2+ query users*) – does not yield a different picture.

Table 1: Average query length and standard deviation

	no-halo	halo
All queries	2.18 (1.36)	2.16 (1.32)
Single query users	2.06 (1.11)	2.15 (1.40)
2+ query users	2.25 (1.48)	2.17 (1.27)

In order to evaluate the suitability of the assumption that longer queries perform better, we compute the block MRR across all the 6,026 queries in our 50-day log (combined over conditions and experiments). Figure 1 shows that in our production system the common assumption that longer queries perform better does not hold: longer queries do not lead to more clicks on better ranked result blocks, and thus do not lead to higher MRR.
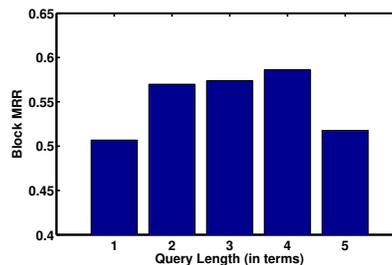
Figure 1: Mean reciprocal rank per query length

4 CONCLUSION

We find that the query halo effect did not entice people to submit longer queries: there was no significant difference between the halo condition and no-halo condition across the two experiments performed. We also find that longer queries, in this context, do not necessarily result in better retrieval performance. The full version of this paper was published at ACM SIGIR 2017 [2].

REFERENCES

[1] Elena Agapie, Gene Golovchinsky, and Pernilla Qvarfordt. 2013. Leading people to longer queries. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3019–3022.
[2] Djoerd Hiemstra, Claudia Hauff, and Leif Azzopardi. 2017. Exploring the Query Halo Effect in Site Search. In *Proceedings of the SIGIR Conference on Research and Development in Information Retrieval*. ACM, 981–984.

---

[1]https://utwente.nl/search
[2]http://searsia.org