

Impact of Search Results on User Queries

Sofia Stamou

CEID, Patras University 26500 GREECE
stamou@ceid.upatras.gr

Lefteris Kozanidis

CEID, Patras University 26500 GREECE
kozanid@ceid.upatras.gr

ABSTRACT

In this paper, we experimentally study how web searchers select the keywords to describe their information needs and specifically we investigate whether query keyword selections are influenced by the results the users reviewed for a previous search. For our study, we determine two types of searches: (i) those in which users define their queries without any external influence and which we call *tightly-focused* and (ii) those in which users define their queries under some external influence and which we call *loosely-focused*. Based on the analysis of the user querying trends and web visits on the query results, we propose a model that tries to capture the results' influence on the specification of the subsequent user queries. The application of our model on a search trace of 19,250 queries issued to Google by 18 users over a period of two months reveals that in overall search results influence the specification of 12.79% of the web queries.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: search process

General Terms

Measurement, Experimentation, Human Factors

Keywords

Web search, search results influence, querying trends

1. INTRODUCTION

Search engines play an important role in our daily activities since their primary objective is to help us find the information we need every time we rely on them. In this paper, we experimentally study how web searchers specify their queries and specifically we investigate whether their query specifications are influenced by the contents of the previously visited search results. Although there exist several works on how people search for information on the web [6][3][10][8], most of the reported works concentrate on identifying the search goals associated with web queries and are less concerned about the impact that the previously visited results might have on query specifications. Thus, our working hypothesis is that existing query classification schemes [1][9] are useful for identifying the different goals associated with web searches but they are less informative about how these goals are specified in terms of search keywords. With that in mind and motivated by the desire to understand the behavioral aspects associated with query verbalization, we define the following query specification types that capture how web searchers establish their querying behavior.

Tightly-focused searches: a user engages in a tightly-focused search when the keywords she selects for verbalizing an information need are self-defined and not influenced by the contents of the previously seen search results.

Loosely-focused searches: a user engages in a loosely-focused

search when the keywords she selects for verbalizing an information need are influenced by the contents of the pages visited for a previous search.

Based on the above search types, we comparatively analyze the user visits on the search results in order to firstly identify the search type that every query represents and then estimate whether the query submission was influenced by the results visited for a previous search. In Section 2, we describe a human survey in which we asked 18 users indicate the criteria under which they make their query selections. The purpose of the survey is to: (i) identify the different query strategies that web users employ and (ii) assess whether and how search results affect the user querying trends. Our findings indicate that there are different criteria under which web searchers make their query specifications and that search results have some influence on information seekers while specifying their query keywords. Based on our survey results, in Section 3, we propose a model for capturing the results' influence on the specification of the subsequent user queries. In Section 4, we present an experiment in which we applied our model for automatically identifying among a set of 19K web searches those that are specified under the direct influence of the previously seen results. Our study shows that the overall results' influence on the specification of search queries amounts to 12.79%.

2. STUDY OF USER QUERYING TRENDS

In this section, we present a human survey we carried out in order to (i) identify the strategies that web searchers employ when picking their query keywords and (ii) formulate a benchmark query set against which we can assess the search patterns that characterize the user querying strategies. Our benchmark set consists of 368 queries issued to Google from students in our department in a single session. To gather the queries, we installed a browser plug-in at the workstations of 18 volunteers from our department and we recorded the headers of their HTTP requests from which we could obtain Google queries and their subsequent clickthrough data. The fact that our participants were competent web searchers does not impose any bias to our findings, since our experimental queries (cf. Section 4) were also collected from the same users. For every query in our test set, we presented to our participants the following selection criteria and asked them to indicate for each of their queries the criterion that best described their query keywords' specification. We instructed our participants to indicate their query selection criteria right after they had issued a query and reviewed the query results.

1. I intended to submit this query when I started searching. The specification of my query keywords was not influenced by the search results I have previously viewed.
2. I did not intend to submit this query when I started searching. The specification of my query keywords was influenced by something I read/heard before issuing the query but not by the results I have reviewed for a previous search.
3. I did not intend to submit this query when I started searching. The specification of my query keywords was entirely influenced by the results I have reviewed for a previous search.
4. None of the above.

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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WIDM'09, November 2, 2009, Hong Kong, China.

Copyright 2009 ACM 978-1-60558-808-7/09/11...\$10.00.

Under the above query selection criteria, the first choice represents tightly-focused searches, while the second and third choices represent loosely-focused searches. Additionally, the third choice represents searches specified under the direct influence of the previously visited results, whereas the second choice represents searches that were externally influenced but not by the previously visited search results. The fourth choice represents queries the specification criteria of which cannot be clearly determined and as such we characterize them as searches of unknown influence.

2.1 Survey Results

Based on the choices indicated for each of the 368 searches that our subjects performed during their participation in the survey, we manually classified our test queries into their corresponding types and we estimated the fraction of the queries that pertain to every type. Table 1 reports obtained results. Note that our test queries adhere to informational search tasks as determined in [3].

Table 1. Distribution of test queries across search types

Queries examined	368
% tightly-focused queries	63.28%
% loosely-focused queries without result influence	22.05%
% loosely-focused queries with result influence	13.85%
% of queries with unknown influence	0.82%

As the table suggests, if we consider the 368 queries as a whole, most of them are associated with a tightly-focused search intention, which implies that our participants in most of their searches specified their query keywords without any external influence. On the other hand, 22.05% of the queries although they had a loose search focus, their verbalization was not influenced by the contents of the previously seen search results, whereas 13.85% of the queries were issued under the direct impact of the previously visited results. To assess how the results' influence is pronounced in the query specifications, we associated every query with its corresponding click-through activity on the query results and organized them in four clusters; one per query specification criterion. Then, for the queries and follow-up clicks that had been assigned to the first three clusters (i.e. tightly-focused, loosely-focused without results' influence and loosely-focused with results' influence), we examined the following features: number of terms per query, amount of result clicks per query and elapsed time between the query submission and the first click on some result. Based on the values computed for the above features for every query, we aggregate values across all queries in a cluster for quantifying the characteristics of the search types examined. Table 2 reports our computations for all the queries assigned to each of the clusters.

Table 2. Features of queries and follow-up clickthrough

Tightly-focused queries		233
Avg. terms/query		2.53
Avg. # of clicks/query		6.9
Avg. time between query submission and first click		1.9 sec
Loosely-focused queries (with no result influence)		81
Avg. terms/query		2.71
Avg. # of clicks/query		7.1
Avg. time between query submission and first click		3.2 sec
Loosely-focused queries (with result influence)		51
Avg. terms/query		3.09
Avg. # of clicks/query		8.2
Avg. time between query submission and first click		4.8 sec

Results show that loosely focused queries contain slightly more terms and yield more result clicks than tightly focused ones. More-

over, the time interval between issuing a query and clicking on a result is longer for loosely focused searches.

To identify how the results' influence is demonstrated in the specification of a follow-up query, we examined the correlation between the query keywords and the contents of the pages visited right before the query submission. In this respect, we downloaded the contents of the pages our users had visited right before the submission of a query, we parsed them to remove markup and we applied string matching to their textual contents in order to compute the terminological overlap between the documents a user visited right before the submission of a query and the terms comprising that query. The terminological overlap between a query q and the contents of the pages P the user visited before issuing q , is given by:

$$Overlap(q, p) = \frac{1}{|P|} \sum_{\substack{q=1 \\ q(t) \in q}}^{|P|} \frac{|q(t) \cap p(t)|}{|q(t)|}$$

Where $q(t)$ denotes the query terms, $p(t)$ denotes the terms in a page p the user has visited for the query issued before q and P is the set of all the pages the user visited for the query preceding the submission of q . Note that under our model, the first user query in a session as well as queries succeeding searches without result clicks are deemed as queries not specified under the results' direct influence, since there are no pages visited prior to their submission. Having computed the terminological overlap between our test queries and their previously visited results, we organized them into three groups: in the first group we assigned the queries that had no term in common with the contents of the previously seen results (*zero overlap*), in the second group we assigned the queries that had all their terms in common with the contents of the previously seen pages (*perfect overlap*) and in the third group we assigned the remaining queries (*partial overlap*). Finally, we associated every query with the search type it represents as indicated by our subjects, in order to examine how the results' impact is pronounced in the keyword specification of the queries that follow them. Table 3 reports the fraction of queries in each of the examined search types that exhibit perfect, partial or no terminological overlap with the contents of the previously seen results.

Table 3. Terminological overlap between queries and results

Searches of no terminological overlap with previous results		205
% of tightly focused queries		89.76%
% of loosely focused queries with no result influence		8.78%
% of loosely focused queries with result influence		1.46%
Searches of perfect terminological overlap with previous results		48
% of tightly focused queries		16.67%
% of loosely focused queries with no result influence		22.92%
% of loosely focused queries with result influence		60.41%
Searches of partial terminological overlap with previous results		112
% of tightly focused queries		36.61%
% of loosely focused queries with no result influence		46.43%
% of loosely focused queries with result influence		16.96%

Results indicate that loosely focused searches specified under the direct impact of the previously visited results exhibit substantial terminological overlap (most of the times perfect terminological match) with the contents of the previously visited results.

3. RESULT'S IMPACT ON USER QUERIES

Based on our survey results, we propose a model that takes as input a user's search logs and performs a number of computations in order to determine which user queries (if any) have been specified under the direct influence of the previously visited results. The search features upon which our model operates are: (i) average number of

terms in the user queries, i.e. $Avg.Q(t)$, (ii) average number of user clicks on the query results, i.e. $Avg.c(Q)$ and (iii) average time interval between the submission of a query and the first user click on a result, i.e. $Avg.t(Q,c)$. Specifically, our model processes a user’s search trace and quantifies the features that characterize the user’s querying and web visitation patterns as shown in Figure 1.

Input: set of user queries Q , number of result clicks per query $c(q)$, elapsed time between query submission and first click on results $t(q,c)$.

Output: user query and web visitation features

*/*Quantify the features that characterize a user’s query and clickthrough trends*/*
For all user queries Q **do**
*/*Estimate average number of query terms*/*
 Compute $Avg.Q(t) = \frac{1}{|Q|} \sum_{\substack{q_i=1 \\ q_i \in Q}}^{|Q|} |q_1(t) + q_2(t) + \dots + q_n(t)|$
*/*Estimate average number of clicks per query*/*
 Compute $Avg.c(Q) = \frac{1}{|C|} \sum_{\substack{q_i=1 \\ q_i \in Q}}^{|C|} |c(q_1) + c(q_2) + \dots + c(q_n)|$
*/*Estimate average time between issuing q and first click*/*
 Compute $Avg.t(Q,c) = \frac{1}{|C|} \sum_{\substack{q_i=1 \\ q_i \in Q}}^{|C|} |t(q_1,c) + t(q_2,c) + \dots + t(q_n,c)|$

Return average scores for query and clickthrough features

Figure 1. Quantifying the query and clickthrough features.

Based on the above computations, our model proceeds with the identification of the user queries that have been specified under the direct impact of the results visited for a previous search. In this respect, our model relies on both the computed user search features and the terminological overlap between the terms in a user query and the terms in the pages visited for a search conducted right before issuing the query. Figure 2 illustrates the steps our model follows for determining which of the queries have been specified under the influence of the previously visited results. Note that $Order(q_i)$ indicates the submission sequence of q_i in the search session so that the $Order(q_i)=1$ implies that q_i is the first query in the session. Moreover, $c(q_{i-1})$ indicates the number of clicks on the results visited for a query preceding q_i , $NRI(q)$ denotes a query q issued under no influence of the previously visited results and $RI(q)$ denotes a query q issued under the direct influence of the previously visited results.

Specifically, the queries our model identifies as being specified under the influence of the previously visited results are those that exhibit all the following features: (i) there is a perfect terminological overlap between the query terms and the contents of the previously visited results, (ii) they contain more terms than the average number of terms in that user queries, (iii) they yield more result clicks than the average user clicks on the query results and (iv) the time interval between issuing the query and visiting a result is above the average amount of time the user needs for clicking on search results. Given the limitation imposed by the requirement that all the above features must apply in order for our model to determine the queries that have been specified under the influence of the previously visited results, our model may misclassify queries that are influenced by the previous results as non-influenced. However, considering that it is impossible to encapsulate in a single model the complex mental processes associated with the query specification task, we decided to build a model that characterizes a query as results’ influenced only when there is solid evidence to support this characterization.

Input: set of user queries Q , scores for the user query and web visitation features, overlap between query and previously visited results $Overlap(q,p)$

Output: Queries specified under the influence of the previously visited results

/ Capture result influence on the specification of queries */*
For each $q_i \in Q$ **do**
 Compute order of q_i in Q
If $Order(q_i) = 1$ **Return** q_i is Non-Results-Influenced, i.e. $q_i \in NRI(Q)$
Else
 Compute clickthrough for the query issued before q_i , i.e. q_{i-1}
If $c(q_{i-1}) = 0$ **Return** q_i is Non-Results-Influenced, i.e. $q_i \in NRI(Q)$
Else
For every $q_i \in NRI(Q)$ **do**
 Extract the textual contents of the pages visited for q_{i-1}
*/*Estimate overlap between $q_i(t)$ and terms in pages $p(t)$ visited for q_{i-1} */*
 Compute $Overlap(q_i,p)$
If $Overlap(q_i,p) = 1$ **and** $q_i(t) \geq Avg.Q(t)$ **and** $c(q_i) \geq Avg.c(Q)$ **and** $t(q_i,c) \geq Avg.t(Q,c)$
Return q_i is influenced from Results of q_{i-1} i.e. $q_i \in RI(Q)$
Else
If $Overlap(q_i,p) \leq 0$ **and** $q_i(t) \leq Avg.Q(t)$ **and** $c(q_i) \leq Avg.c(Q)$ **and** $t(q_i,c) \leq Avg.t(Q,c)$
Return q_i is Non-Results-Influenced, i.e. $q_i \in NRI(Q)$
Else
Return q_i may be influenced

Return query influence

Figure 2. Identifying the results’ influenced queries.

4. EXPERIMENTAL EVALUATION

In this section we present an experiment we carried out in which we applied our model on a search trace in order to estimate the fraction of queries that are specified under the direct influence of the previously visited results’ contents. For our experiment, we collected the complete search trace issued to Google by the 18 volunteers who participated in our survey over a period of two months. Due to space constraints, we do not delve into the details of processing and cleaning the collected search traces but rather report the statistics of our dataset (cf. Table 4).

Table 4. Statistics on the experimental dataset

Collection period	May– June 2009
# of unique user ID’s	18
# of queries	19,250
# of unique queries	9,111
# of clickthrough events	102,795
# of queries w/o clickthrough	2,675
# of sessions	5,034
avg. # of sessions/user	279.7
avg. # of queries/session	3.8
avg. # of clicks/query w clickthrough	6.2

Based on our collected search trace, we grouped the recorded user queries and clickthrough data into individual sessions. To identify the user queries issued under the direct influence of the previously visited results, we employed our model and firstly estimated the features that characterize every user search trends (cf. computations in Figure 1). Then, we explored the contents of the pages visited prior to a query submission, and based on the computations reported in Figure 2 we determined which queries were specified under the direct impact of the previously visited search results. Table 5 reports our experimental results.

Table 5. Results' impact on the specification of search queries

Experimental Queries	19,250
% queries specified under no results influence	63.64%
% queries specified under the results direct influence	12.79%
% of queries the impact of which needs manual validation	23.57%

As the table shows, 12.79% of the experimental queries were identified by our model as specified under the direct influence of the previously seen results' contents. This implies that those queries had all their terms in common with the contents of the previously visited results and they also contained more terms than the average number of terms in that user queries. In addition, those searches resulted in more page visitations and required more time before a first page visit compared to the average number of clicks and time before the first click that characterize the search trends of the user who issued those queries. Moreover, our findings indicate that for 23.57% of the examined queries, our model could not make a definite decision about whether these have been specified under the previous results' impact or not. This means that those queries had some terminological overlap (it could be perfect overlap in some cases) with the contents of the previously visited results but some or all of their feature estimations scored below the average estimations that characterize the search trends of the users who issued those queries. Considering the restrictions we have applied to our model and which we justified earlier, we assume that if we relaxed the constraints imposed in order for our model to characterize a query as influenced by the contents of the previously visited results, we might get a bit increased numbers with respect to the influenced queries. Nevertheless, as already explained we designed our model in a way that it makes a positive judgment only when there is plenty of evidence to support it.

Based on the results delivered for the queries submitted by our participants and considering that our findings rely on the features detected for characterizing our participants' search trends, we may conclude the following: (i) results generally agree with the findings of our human survey in which the same users determined the specification criteria of their queries explicitly and (ii) results prove the need to investigate the impact of search results on the specification of their following queries in a systematic manner. Thus, although our study is limited to a small fraction of the web population and searches, it provides evidence that search results play a role in our querying patterns; an issue that has not been systematically addressed into a focused evaluation setting. We hope that our findings will help researches build enriched search models by taking into account not only what people are searching for but also how they search.

5. RELATED WORK

There exist several works that rely on the analysis of the user activity on the search results for modeling the searchers' behavior. Researchers have studied the behavioral aspects of the users' search activities [9] and tried to build models that explain such behavior [7]. There has also been substantial work on deriving the characteristics of web queries [4] [3] and on analyzing the users' clicking behavior after they issue a query [2] in order to identify similar web requests. Researches have also investigated the extent to which people's search behaviors differ in terms of interaction styles and search term selection strategies [11]. However, most of the works in

this direction distinguish searchers into domain experts and novices and comparatively analyze their web interactions in order to capture their search behaviors [5] [12]. Although previous studies have shown differences in search queries and strategies between different user groups, they mainly focus on modeling such differences for improving retrieval performance and ultimately for assisting information seekers experience successful web searches. Our study builds upon existing works and investigates how much searchers are influenced by the contents of the visited results in verbalizing their next search requests.

6. CONCLUDING REMARKS

In this paper, we studied how web information seekers pick the search keywords to describe their information needs and specifically we examined whether query keyword specifications are influenced by the results the users reviewed for a previous search. Based on the analysis of the user querying trends and web visits on the query results, we proposed a model that tries to capture the results' influence on the specification of the subsequent user queries. The application of our model on a search trace of 19,250 queries issued to Google by 18 users over a period of two months reveals that in overall search results influence the specification of 12.79% of the web queries. In the future, we plan to carry out a large-scale laboratory study in order to assess the robustness of our results over changing samples.

7. REFERENCES

- [1] Broder, A. 2002. A taxonomy of web search. In SIGIR Forum, 36(2).
- [2] Davison, B.D., Deschenes, D.G. and Lewanda, D.B. 2003. Finding relevant web site queries. In Proc. of the 12th WWW Conference.
- [3] Jansen, B.J., Booth, D.L. and Spink, A. 2008. Determining the informational, navigational and transactional intent of web queries. *Information Processing & Management* 44:1251-1266.
- [4] Jansen, B.J., Spink, A. and Saracevic, T. 2000. Real life, real users and real needs: a study and analysis of user queries on the web. In *Information Processing and management*, 36(2): 207-227.
- [5] Hembrooke, H.A., Gay, G.K. and Granka L.A. 2005. The effects of expertise and feedback on search term selection and subsequent learning. In *JASIST*, 56(8): 861-871.
- [6] Kammenhuber, N., Luxenburger, J., Feldmann, A. and Weikum, G. 2006. Web search clickstreams. In Proc. of the 6th ACM SIGCOMM Conference on Internet Measurement, pp. 245-250.
- [7] Kang, I. and Kim, G. 2003. Query type classification for web document retrieval. In Proc. of the 26th SIGIR Conf., pp. 64-71.
- [8] Pass, G., Chowdhury, A. and Torgeson, C. 2006. A picture of search. In Proc. of the 1st Conf. on Scalable Information Systems.
- [9] Rose, D. and Levinson, D. 2004. Understanding user goals in web search. In Proc. of the 13th WWW Conference, pp. 13-19.
- [10] Spink, A., Park, M., Jansen, B. and Pedersen, O. 2006. Multitasking during web search session. *Information Processing and Management*, 42(5): 1366-1378.
- [11] White, R. and Drucker, S. 2007. Investigating behavioral variability in web search. In Proc. of the 16th WWW Conference, pp. 21-30.
- [12] White, R., Dumais, S. and Teevan, J. 2009. Characterizing the influence of domain expertise on web search behavior. In Proc. of the 2nd Intl Conference on Web Search and Data Mining, pp. 132-141.