Do Topic Shift and Query Reformulation Patterns Correlate in Academic Search?

Xinyi Li and Maarten de Rijke

University of Amsterdam, Amsterdam, The Netherlands
{x.li,derijke}@uva.nl

Abstract. While it is known that academic searchers differ from typical web searchers, little is known about the search behavior of academic searchers over longer periods of time. In this study we take a look at academic searchers through a large-scale log analysis on a major academic search engine. We focus on two aspects: *query reformulation patterns* and *topic shifts in queries*. We first analyze how each of these aspects evolve over time. We identify important query reformulation patterns: revisiting and issuing new queries tend to happen more often over time. We also find that there are two distinct types of users: one type of users becomes increasingly focused on the topics they search for as time goes by, and the other becomes increasingly diversifying. After analyzing these two aspects separately, we investigate whether, and to which degree, there is a correlation between topic shifts and query reformulations. Surprisingly, users' preferences of query reformulations correlate little with their topic shift tendency. However, certain reformulations may help predict the magnitude of the topic shift that happens in the immediate next timespan. Our results shed light on academic searchers' information seeking behavior and may benefit search personalization.

1 Introduction

Academic search deals with the retrieval of information resources in the domain of scientific literature. Hemminger et al. [15] point out that academic search engines have become the primary portal for researchers to gain information; see also [31]. In recent years, there have been several publications focused on academic search and academic searchers. However, most are very limited in scale, and rarely reveal insights into the search behavior of academic searchers based on the analysis of large-scale transaction logs [14, 23, 24]. In this study we take a look at academic search through a large-scale log analysis from a major academic search engine.

Academic searchers do have a distinct search pattern that is different from the typical web searchers. For instance, in web search, the search activity becomes the least intensive on Fridays and peaks in the weekends [2]. But, as shown in Fig. 1, academic search activity peaks during weekdays, and drops in the weekends.

To study the behavior of academic searchers, we investigate two key aspects: *query reformulations* and *topic shifts*. Both have received much attention in user behavior studies of web search [4, 18, 26], but to the best of our knowledge, there is no previous work on revealing the query reformulation behavior and topic shifts of academic searchers that is based on a large-scale log analysis. In fact, very little is known about these two aspects of academic search.

Through this study, we provide answers to 3 research questions:

RQ1 What is the query reformulation behavior of academic searchers?

RQ2 Do academic searchers have shifts in topical interests over time?

RQ3 Is there a correlation between query reformulation behavior and topic shift?

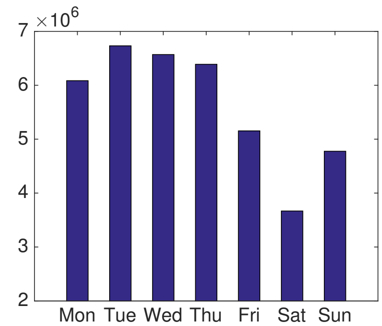


Fig. 1: Average number of queries per weekday in academic search (based on the dataset described in §3).

For the first question, we look at query reformulation behavior over time. Query reformulation happens after the user has examined the search engine result page and provides a more explicit type of feedback than clicks, which are implicit and noisy [9]. We look at five frequent types of query reformulation: *revisiting a previous query*, *adding terms*, *dropping terms*, *substituting part of the query*, and *issuing a completely new query*. We study how the type of reformulation behavior changes over time and find that revisiting and issuing new queries tend to happen more often as search goes on.

For the second question we take a quantitative approach to study topic shift over time. We train an LDA model [3] on all long sessions in the query log that we examine. We segment a user's queries into different timespans, and treat queries in each timespan as a bag of words. We infer a topic vector for each timespan of the user. Topic shift between successive timespans is then calculated using the Euclidean distance between the topic vectors. In this process we identify two types of user: one type increasingly focuses on topics over time and the other diversifies over time.

Finally, we conduct a correlation study to see how these two aspects—query reformulation and topical shift—are correlated with each other. We find that user's query reformulation patterns have little correlation with the tendency of topic shift, meaning that users with distinct reformulation preferences in search could be equally likely to be diversifying or focusing on topics. We also find that certain reformulations (viz. *adding terms* and *issuing new queries*) may help predict the magnitude of the next topic shift.

Contrary to previous work that studies academic searchers through surveys and user studies, this paper sheds light on the reformulation behavior and topical shifts of academic searchers through a large-scale log analysis. The insights gained help us to understand academic searchers' information seeking patterns from a much larger user base, and may be useful for personalization in academic search.

In §2 we discuss related work. In §3 we introduce the dataset characteristics. In §4 we describe our approach to study query reformulations and topic shifts. In §5 we show the result and analysis from the correlation studies. We present our conclusions in §6.

2 Related work

2.1 Academic search Academic search involves the indexing and retrieval of information objects (papers, journals, authors, ...) in the domain of academic research. The earliest academic search engine MEDLINE, which began functioning in 1971, allowed a maximum of 25 simultaneous users [28]. It was restricted to library usage and only pre-programmed searches were supported instead of online queries.

When the web became popular in the 1990s, online academic search engines started to flourish and gained popularity. Typical examples are Citeseer [11] and Aminer [38], which focus on metadata retrieval and academic network extraction respectively. There are several surveys and user studies on the search behavior of researchers on modern academic search engines [31–33], which are based on a relatively small sample of researchers. The few log analyses conducted on search engines of digital libraries are either investigating a single discipline [14, 23], or limited in scale [24], as a result of which they are not representative of academic search. Moreover, they focus on basic usage statistics and lack insights on user behavior in search sessions. Recently, Li et al. [27] studied the user behavior and query failure phenomenon in academic search through a large-scale transaction log analysis.

2.2 Query reformulations Query reformulation is an important aspect of user behavior during search sessions. In recent years, there has been a range of studies that cover patterns and models of query reformulation [4, 6, 18, 25, 35, 36], how they work in a collaborative setting [30], in voice search [20] or in mobile search [37], and their applications [5, 7, 21, 34]. These studies show that query reformulations are the key to understanding user behavior, which will benefit retrieval tasks such as query auto completion [21] as well as topic and intent finding in users' queries [34], and which may help improve retrieval performance [13]. The findings are mostly in the domain of web search and the query reformulation behavior studied is that of the general web users.

Multiple category schemes have been used for query reformulation in the literature [4, 6, 18, 25, 35, 36]. Different category schemes may correspond to (1) search engines of different designs (e.g., whether searches on multiple verticals are supported), (2) whether using search assistance is considered as a reformulation such as query suggestion, or (3) different granularities of query reformulations. Manual categorization may provide fine-grained results [6, 25, 35] but can not easily scale up to large query logs. On the other hand, rule-based [18, 36] or learning-based [4] methods can be applied to a large query log, and are thus more suitable for analyzing long term query reformulations from a large user base.

2.3 Topic shift in queries There has been a whole line of research that investigates topic mining in web search query logs [1, 16, 17, 22], where the emphasis is on how to segment and cluster queries by topic. However, the multi-tasking nature of web searchers, which means searching and switching between multiple topics within and across sessions [29], makes it cumbersome to derive useful insights from users' topic shifts, especially over long periods.

This paper differs from previous work in academic search, by studying a large transaction log from a major academic search engine, with a focus on user behavior in search sessions. The findings are therefore better able to represent academic searchers, compared with earlier small-scale user studies and surveys. It also differs from previous work in query reformulations in web search, by revealing the academic searchers' preferences instead of those of the general web users. The paper differs from work on topic shifts in web search by looking at a different domain: academic search. Compared to the web searchers who have diverse, parallel, and fast-shifting topic interests, academic searchers are more likely to have consistent interests in a general topic. For instance, a researcher in information retrieval is more likely to stay in this general topic than diverting to biology sciences. This makes studying the long term topic shift pattern meaningful. Moreover, this study tries to link query reformulation to topic shift, and provides useful insights into their connections through a series of correlation studies.

3 Data

We study a query log from the ScienceDirect search engine,[1] containing over 39 million queries. The query log is collected from September 28, 2014 to March 5, 2015. Table 1 shows the length statistics of the query log. Two thirds of the traffic come from institution-authorized access, meaning that users in a certain IP range can access the search engine, and they share the same session ID and user ID in the query log. Besides, many institutions use proxies or firewalls so that their IP is recorded instead of the terminal device. Therefore it is not possible to differentiate these IP-users. We are only confident in an ID-user one-to-one mapping when they log in or access the search engine from outside the institution. And we study these "non-IP" users only, who contribute about one third of the traffic.

Table 1: Query length statistics in word count.

Category	#N	min	max	mean	median
Sciencedirect	39M	1	419	3.77	3

With a timeout of thirty minutes as a threshold, there are a total of 4,307,889 sessions for these non-IP users, and 2,833,549 of them contain at least 3 queries which we denote as "long sessions." To obtain enough data of users, we confine the scope of users to those who have a minimum of 30 queries, and whose search behavior lasts over 30 days at least. This leaves us with 29,093 users and 1,918,334 query records.

4 Approach

In this section we describe how we study the behavior and topic change of academic searchers in a series of correlation studies.

First, we highlight the statistics of the prominent types of query reformulations from the query log. Then, we apply a time sequence-based method to make observations of how users progress in search. We break each user's queries into sequences and then align them, so that we can compare how users progress during search even if they start at a different time. Specifically, we put each user's queries into bins separated by a certain length of timespan (to be specified below). Then, we align all searchers' queries by timespan, with the first timespan of a user denoted as 0, the second as 1, in a natural number sequence. We can observe query reformulation and topic shift of users as they move from one timespan to the next. In this case, to gain enough samples from the dataset and also to ensure statistical significance in our later correlation analyses, we sample timespans of 3, 7 and 14 days long. We choose timespans of different lengths to observe whether some changes are more prominent over longer timespans.

[1] http://sciencedirect.com

The length of timespans chosen also corresponds with the usual information seeking cycles of academic searchers, as research suggests that information-seeking happens toward a weekly basis rather than daily basis for faculty and graduate students [8, 31]. Note that users may issue no query in a certain timespan; in such cases the timespan will be neglected for that user.

Query reformulation tendency over time. To uncover the reformulation preferences for the academic searchers as a whole, we examine the query reformulation preference over time for all academic searchers combined. For each timespan, we aggregate the frequency of each reformulation from all users and obtain the proportion of each reformulation type. We hypothesize that certain reformulations might happen more frequently as time goes on, for instance revisiting, because academic searchers tend to have a consistent interest in their field of study [19] and may thus need to submit a previous query repeatedly in search of new information. We try to determine if there is indeed a linear correlation of the proportion of an action over the course of time (represented as a natural number sequence of timespans). To this end, we use Pearson's correlation.

It is common for users to use a combination of the query reformulations listed in the previous section (revisiting, adding a term, dropping a term, substituting a term, new query) in order to reach their search goal. In our analysis, we calculate the proportion of each query reformulation in each time span for every user.

Topic shift. We study the tendency of a user to shift topic over time with a quantitative approach as we aim to measure the magnitude of change in topic. We train an LDA model on long sessions that contain at least 3 queries. Each session is treated as a "document" in training because the queries within a single session mostly likely belong to the same general topic. The number of topics is set to 150, which is a reasonable value in the academic domain [12] and also ensures relatively fast convergence in Gibbs sampling. For each user, we model the queries in each timespan as a bag of words and use the trained LDA model to infer a topic vector. Then, for a given user the magnitude of topic shift between adjacent timespans is calculated using the Euclidean distance between the user's topic vectors for the two timespans.

Correlations. After studying how users' reformulation behavior and topical interest change over time, respectively, we aim to find whether there is a correlation between a user's query reformulation patterns and their topic shift tendency. Specifically, we look at two aspects of the correlations. First, the macroscopic aspect, i.e., whether a user's topic shift tendency is correlated with query reformulation preferences. For instance, suppose a user favors a specific type of reformulation, say substitution; is this user likely to be diversifying in topic shifts? Second, there is the microscopic aspect: in successive timespans, is the proportion of each reformulation type in the first timespan correlated with the topic change that happens during the next timespan? Based on the correlation findings, we consider the task of predicting the magnitude of a user's topic shift during the next timespan.

5 Results and analysis

In this section we present the results of our analysis of users' query reformulations and topic shifts. We first analyze these two aspects separately and then perform a series of correlation studies to examine their connections.

5.1 Query reformulation types To study users' query reformulation types, we apply a syntactic-based automatic categorization. Our taxonomy does not require human annotations and does not have the fine-granularity of those methods in [4, 18, 36]. However it is fully unsupervised and is scalable to a large query log; it contains five reformulation types that are common to the majority of taxonomies previously used for query reformulations [4, 6, 18, 25, 35, 36]. The main difference is that none of these previous publications considers "revisiting queries" as a reformulation while we do (Bruza and Dennis [6] consider "repeated query" but there is no user identifier in their query log).

Revisiting Revisiting is issuing a query that is already in the user's search history [39]. In academic search, we find that this reformulation type is very prominent, making up 33.8% of all reformulations, which shows that academic searchers tend to have some consistent search intents and will seek information on the same topic repeatedly.

Adding terms This type of reformulation is characterized by adding at least one term to the previous query, and corresponds to the process of refining search. This is typically seen in sessions where users start with a general query on a certain topic, then add terms to examine sub-aspects within the topic [35]. This reformulation type constitutes 8.5% of all reformulations.

Dropping terms This is the opposite process of the adding reformulation type, constituting 5.6% of all reformulations. By dropping at least one term from the previous query, the user aims to retrieve information that is more general than the previous query [35]. This may happen when academic searchers need context information during learning.

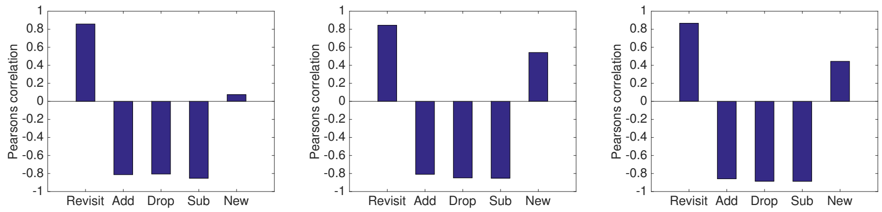
Substituting terms Substitution of terms is the second most prominent reformulation type that accounts for 28.0% of all reformulations. Substitution means keeping certain at least one term in the original query intact, then dropping old terms and adding new terms. Substitution behavior may happen when a user is refining a search, e.g., changing a synonym, or when the user is exploring different aspects about a certain topic [35].

New query This reformulation concerns the situation where the user Issues a query that has no overlap of words with the previous query and that does not appear in the user's search history. Submitting a new query that is different often means a change of search intent [4]. It happens when other reformulations will not address the new intent of the users. New queries make up 24.1% of all reformulations.

Compared to web search, where substituting terms accounts for the most popular type (ranging from 22.73% to 37.5% in different datasets [4, 18]), the most prominent type in academic search is revisiting and substituting terms only comes next.

5.2 Query reformulation tendency for all academic searchers combined Fig. 2 plots all searchers' query reformulation tendency.

By definition of the correlation strength [10], there is a "very strong" positive correlation of the proportion of revisiting behavior over time, in the analyses of all timespans. This confirms our earlier hypothesis in §4, that there is an increasing trend of revisiting queries by academic searchers, which shows their consistent interests in certain topics. Interestingly, between timespans of 3 days, the tendency to submit new queries is weak, but at longer timespans (7 or 14 days), we can observe a moderate positive correlation. This suggests that submitting new queries tends to happen not immediately (within a 3

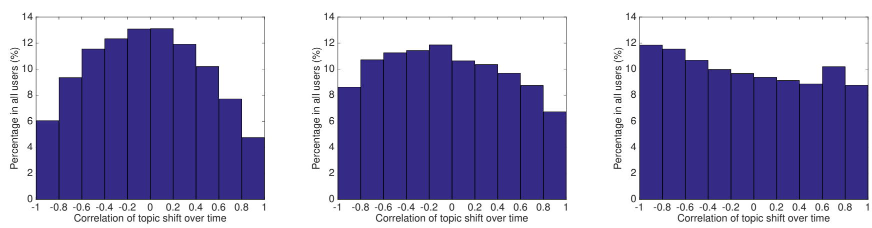


(a) Timespan = 3 days. **(b) Timespan = 7 days.** **(c) Timespan = 14 days.**

Fig. 2: The query reformulation preference over time for all the academic searchers, measured in correlation of the proportion of the reformulation actions (revisiting, adding terms, dropping terms, substitution and new query) over time.

day gap), but within a longer gap. The negative correlation for the other three reformulations (add, drop, and substitute) shows that users perform these reformulations less frequently in the later period of search.

5.3 Topic change tendency Using the approach described in Section 4, we study the magnitude of the users' topic shift over time. The tendency is represented by the correlation strength: the larger the correlation, the bigger the topic shift over time for a user. Fig. 3 shows the distribution of the correlation of the users, for 3 different timespans.



(a) Timespan = 3 days **(b) Timespan = 7 days.** **(c) Timespan = 14 days.**

Fig. 3: The correlation of user topic shift over time.

The correlation strength of topic shift over time indicates the evolution of user interests over time, namely whether they tend to become more focused or more diversified. In general, we find that nearly half the users tend to have increasing topic shifts over time (diversifying), and the other half have decreasing shifts (focusing). For different timespans, we see from the shape of the distribution, that there are more users showing a stronger tendency of topic shift (either positive or negative) as the timespan increases. This indicates that bigger topic shifts tend to happen when the time gap between searches is longer.

5.4 Correlation between reformulation behavior and topic shift There are users who become more focused over time and those who do not. Correspondingly, we group users by their tendency to shift topics, and study if this tendency has a correlation with query reformulation patterns. Specifically, users are divided into 6 groups by the Pearson correlation strength r of the topic shift tendency over time: moderately diversifying ($0.4 \leq r < 0.6$), strongly diversifying ($0.6 \leq r < 0.8$), very strongly diversifying ($0.8 \leq r \leq 1.0$) and moderately focused ($-0.6 \leq r < -0.4$), strongly focused ($-0.8 \leq r < -0.6$), very strongly focused ($-1.0 \leq r < -0.8$). Then we look at the correlation with the user's different reformulation type's proportions, as shown in Fig. 4.

Fig. 4 shows that we cannot differentiate diversifying or focused users, purely based on their query reformulation patterns. That is, the user's preference of choosing certain query reformulations is not correlated with their topic shift tendency. This is an interesting finding as it shows that even users with distinct query reformulation preferences, could be equally likely to be focusing or diversifying in search.

Taking a step back, although we cannot determine whether a user is focusing or diversifying based on preference of reformulations, can we predict the magnitude of topic shift to happen in the next timespan given only the user's current reformulation behavior? To answer this question we first examine the individual correlation between the proportion of each reformulation type at a given timespan, with the topic change that happens at the next timespan. See Table 2.

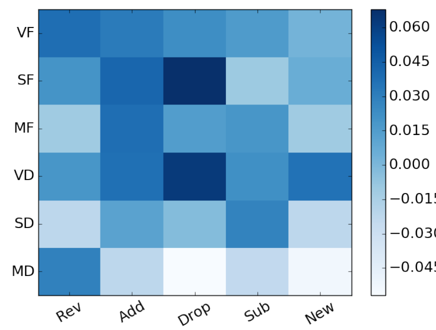


Fig. 4: The correlation of the topic shift tendency (MD: moderately diversifying, SD: strongly diversifying, VD: very strongly diversifying, MF: moderately focused, SF: strongly focused, VF: very strongly focused), with the proportion of the reformulation actions (revisiting, adding terms, dropping terms, substitution and new query) for each user.

Individually, for the majority of users there is only a weak correlation between a query reformulation type and the next topic shift. For users who show a strong correlation ($-1.00 \leq r < -0.60$ or $0.60 \leq r \leq 1.00$), submitting new queries contributes the least to a decrease in topic shift magnitude and also the most to an increase in topic shift magnitude, respectively, compared with other reformulation types. For longer timespans, there are more users who exhibit a strong correlation. Especially when the timespan is 14 days, 21.0% of the users show a strong or very strong correlation between adding terms and topic change, and the number is even higher at 23.7% for submitting new queries. Interestingly, substituting reformulations tend to correlate the least with topic change. This suggests that users tend to stay in the same general topic, or a subtopic within the general topic, while modifying only part of the original queries.

5.5 Predicting the magnitude of the next topic shift Next, we try to utilize the observational insights that we have just gained for a prediction task: can query reformulation signals help to predict the magnitude of a user's topic shift?

More precisely, we use features from users' reformulations to predict the magnitude of topic shift at the next timespan. The features are the proportions and number of occurrences of query reformulations in a timespan. We cast this task as a regression task. Our training set is comprised of pairs of query reformulations and the topic shift to happen at the next timespan for all users. The test set consists of the second-last query reformulations and the next (final) topic shift for each user.

Table 2: Correlation of reformulation behavior with topic shift at the next timespan. Each column shows the distribution of users (in percentage) who have different correlation strengths between a reformulation type and topic shift, in an interval of 0.2.

Correlation	Revisit	Add	Drop	Sub	New
Timespan = 3 days					
$[-1.00, -0.80]$	1.6%	1.4%	2.4%	3.7%	0.9%
$[-0.80, -0.60]$	3.7%	2.9%	4.3%	8.1%	2.3%
$[-0.60, -0.40]$	6.9%	6.4%	7.7%	12.9%	4.5%
$[-0.40, -0.20]$	12.7%	11.5%	11.7%	16.9%	7.7%
$[-0.20, 0.00]$	16.5%	15.5%	16.1%	17.2%	11.4%
$[0.00, +0.20]$	17.2%	17.5%	17.6%	14.2%	15.7%
$[+0.20, +0.40]$	15.8%	16.1%	15.1%	11.8%	19.3%
$[+0.40, +0.60]$	13.1%	14.5%	12.9%	7.6%	18.9%
$[+0.60, +0.80]$	8.4%	9.5%	8.0%	4.9%	13.5%
$[+0.80, +1.00]$	4.1%	4.6%	4.2%	2.9%	6.0%
Timespan = 7 days					
$[-1.00, -0.80]$	3.1%	2.4%	3.7%	5.9%	1.7%
$[-0.80, -0.60]$	5.3%	4.5%	6.1%	10.1%	3.4%
$[-0.60, -0.40]$	8.7%	7.4%	8.4%	13.1%	5.7%
$[-0.40, -0.20]$	11.7%	11.7%	11.2%	14.3%	8.4%
$[-0.20, 0.00]$	13.4%	13.9%	13.4%	14.4%	10.9%
$[0.00, +0.20]$	14.7%	14.4%	14.4%	12.7%	13.8%
$[+0.20, +0.40]$	13.7%	13.8%	14.2%	10.6%	16.4%
$[+0.40, +0.60]$	13.1%	14.2%	12.8%	8.6%	17.4%
$[+0.60, +0.80]$	9.9%	10.9%	9.3%	6.4%	13.7%
$[+0.80, +1.00]$	6.3%	6.9%	6.4%	4.0%	8.6%
Timespan = 14 days					
$[-1.00, -0.80]$	4.8%	3.9%	5.5%	8.0%	3.0%
$[-0.80, -0.60]$	7.0%	6.4%	7.1%	11.6%	5.1%
$[-0.60, -0.40]$	8.6%	8.2%	8.9%	12.6%	7.3%
$[-0.40, -0.20]$	11.1%	10.7%	11.1%	12.2%	9.0%
$[-0.20, 0.00]$	11.6%	12.3%	10.9%	11.9%	10.6%
$[0.00, +0.20]$	12.7%	12.3%	11.4%	11.0%	12.1%
$[+0.20, +0.40]$	12.7%	12.5%	12.8%	10.6%	14.5%
$[+0.40, +0.60]$	12.0%	12.8%	12.6%	8.5%	14.7%
$[+0.60, +0.80]$	10.9%	12.2%	11.1%	7.5%	13.5%
$[+0.80, +1.00]$	8.6%	8.8%	8.6%	6.0%	10.2%

We use linear regression and three evaluation measures: correlation coefficient, mean absolute error (MAE) and root mean squared error (RMSE). The prediction results are listed in Table 4. Prediction is more accurate on shorter timespans, with the 3 day predictions reaching a medium correlation ($r = 0.4530$), while 14 day predictions being

Table 3: Query reformulation features for prediction of the magnitude of topic shift at the next timespan.

Name	Description
Reformulation proportions	
Revisiting_Percentage	Percentage of revisiting reformulations
Adding_Percentage	Percentage of adding term reformulations
Dropping_Percentage	Percentage of dropping term reformulations
Substitution_Percentage	Percentage of substituion reformulations
New_Query_Percentage	Percentage of new query reformulations
Reformulation occurrence numbers	
Revisiting_Number	Number of revisiting reformulations
Adding_Number	Number of adding reformulations
Dropping_Number	Number of dropping term reformulations
Substitution_Number	Number of substituion reformulations
New_Query_Number	Number of new query reformulations

at only $r = 0.3225$. The performance difference indicates that topic shift magnitude in a shorter timespan is easier to predict than longer timespans.

Table 4: Linear regression results (correlation coefficient, mean absolute error, root mean squared error) for predicting the magnitude of a topic shift in the next timespan given query reformulation features in the current timespan.

	3 days	7 days	14 days
Correlation Coefficient	0.4530	0.3906	0.3225
MAE	0.0697	0.0755	0.0805
RMSE	0.0931	0.0999	0.1057

6 Conclusion

In this study we have examined users' query reformulation behavior and their tendency of topic shift in academic search through a large-scale log analysis. We have found that over time, academic searchers as a whole tend to conduct revisiting, as well as submitting completely new queries. This pattern corresponds to the academic searcher's information needs: either seeking previous search results or new results on the same search intents, or simply pursuing new search intents. We have identified two types of topic shift patterns in users, namely the focusing type and the diversifying type.

Through a series of correlation studies, we have found that a user's preference for certain query reformulations does not correlate to their topic shift tendency. Nevertheless, users's current reformulation patterns (adding terms, submitting new queries) may help to predict the magnitude of topic change in the immediate next timespan. We further used features from query reformulations for predicting the magnitude of the next topic shift. The findings of the query reformulation behavior, topic shift type, and their connections help to improve our understanding of the behavior of academic searchers from a large user base. They may provide hints for personalized search, such as whether to provide exploratory or focusing type of search results, and recommendations of queries or papers for users.

In future work we intend to look at query reformulation patterns in the context of different search tasks, e.g., a navigational task for a single document, or a learning task for a certain research topic. And we will examine the utility of using query reformulation features to improve retrieval performance and provide better recommendations in academic search.

Acknowledgements. This research was supported by the China Scholarship Council and Elsevier. All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.

Bibliography

[1] L. M. Aiello, D. Donato, U. Ozertem, and F. Menczer. Behavior-driven clustering of queries into topics. In *CIKM*, pages 1373–1382. ACM, 2011.

[2] S. M. Beitzel, E. C. Jensen, A. Chowdhury, O. Frieder, and D. Grossman. Temporal analysis of a very large topically categorized web query log. *J. Assoc. Inform. Sci. & Tech.*, 58(2): 166–178, 2007.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Machine Learning Res.*, 3:993–1022, 2003.

[4] P. Boldi, F. Bonchi, C. Castillo, and S. Vigna. Query reformulation mining: models, patterns, and applications. *Inform. Retr.*, 14(3):257–289, 2011.

[5] I. Bordino, C. Castillo, D. Donato, and A. Gionis. Query similarity by projecting the query-flow graph. In *SIGIR*, pages 515–522. ACM, 2010.

[6] P. D. Bruza and S. Dennis. Query reformulation on the internet: Empirical data and the hyperindex search engine. In *RIAO*, 1997.

[7] F. Cai and M. de Rijke. A survey of query auto completion in information retrieval. *Found. Trends in Inform. Retr.*, 10(4), 2016.

[8] J. Catlow, M. Górny, and R. Lewandowski. Students as users of digital libraries. *Qual. Quant. Meth. in Lib.*, 65:60–1, 2015.

[9] A. Chuklin, I. Markov, and M. de Rijke. *Click Models for Web Search*. Morgan & Claypool Publishers, July 2015.

[10] J. S. B. T. Evans and D. E. Over. *Reasoning and Rationality*. Psychology Press, 1996.

[11] C. L. Giles, K. D. Bollacker, and S. Lawrence. Citeseer: An automatic citation indexing system. In *DL*, pages 89–98. ACM, 1998.

[12] T. L. Griffiths and M. Steyvers. Finding scientific topics. *Proc. Nat. Acad. Sci.*, 101(suppl 1):5228–5235, 2004.

[13] D. Guan, S. Zhang, and H. Yang. Utilizing query change for session search. In *SIGIR*, pages 453–462. ACM, 2013.

[14] H. Han, W. Jeong, and D. Wolfram. Log analysis of academic digital library: user query patterns. *iConference 2014 Proceedings*, pages 1002–1008, 2014.

[15] B. M. Hemminger, D. Lu, K. Vaughan, and S. J. Adams. Information seeking behavior of academic scientists. *J. Assoc. Inform. Sci. & Tech.*, 58(14):2205–2225, 2007.

[16] Y. Hu, Y. Qian, H. Li, D. Jiang, J. Pei, and Q. Zheng. Mining query subtopics from search log data. In *SIGIR*, pages 305–314. ACM, 2012.

[17] W. Hua, Y. Song, H. Wang, and X. Zhou. Identifying users' topical tasks in web search. In *WSDM*, pages 93–102. ACM, 2013.

[18] B. J. Jansen, D. L. Booth, and A. Spink. Patterns of query reformulation during web searching. *J. Assoc. Inform. Sci. & Tech.*, 60(7):1358–1371, 2009.

[19] W. Jeng, D. He, and J. Jiang. User participation in an academic social networking service: A survey of open group users on mendeley. *J. Assoc. Inform. Sci. & Tech.*, 66(5):890–904, 2015.

[20] W. Jeng, J. Jiang, and D. He. Users' perceived difficulties and corresponding reformulation strategies in Google Voice Search. *J. Lib. Inform. Studies*, 14(1):25–39, 2016.

[21] J.-Y. Jiang, Y.-Y. Ke, P.-Y. Chien, and P.-J. Cheng. Learning user reformulation behavior for query auto-completion. In *SIGIR*, pages 445–454. ACM, 2014.

[22] R. Jones and K. L. Klinkner. Beyond the session timeout: automatic hierarchical segmentation of search topics in query logs. In *CIKM*, pages 699–708. ACM, 2008.

[23] S. Jones, S. J. Cunningham, R. McNab, and S. Boddie. A transaction log analysis of a digital library. *Intern. J. Dig. Lib.*, 3(2):152–169, 2000.

[24] H.-R. Ke, R. Kwakkelaar, Y.-M. Tai, and L.-C. Chen. Exploring behavior of e-journal users in science and technology: Transaction log analysis of Elsevier's ScienceDirect onsite in Taiwan. *Lib. & Inform. Sci. Res.*, 24(3):265–291, 2002.

[25] T. Lau and E. Horvitz. Patterns of search: analyzing and modeling web query refinement. In *User Modeling*, pages 119–128. Springer, 1999.

[26] L. Li, H. Deng, Y. He, A. Dong, Y. Chang, and H. Zha. Behavior driven topic transition for search task identification. In *WWW*, pages 555–565. ACM, 2016.

[27] X. Li, R. Schijvenaars, and M. de Rijke. Investigating queries and search failures in academic search. *Inform. Proc. & Manag.*, 2017. To appear.

[28] D. Lindberg. Internet access to the National Library of Medicine. *Effective Clinical Practice*, 3(5):256–260, 2000.

[29] R. Mehrotra, P. Bhattacharya, and E. Yilmaz. Uncovering task based behavioral heterogeneities in online search behavior. In *SIGIR*, pages 1049–1052. ACM, 2016.

[30] A. S. Mohammad Arif, J. T. Du, and I. Lee. Examining collaborative query reformulation: a case of travel information searching. In *SIGIR*, pages 875–878. ACM, 2014.

[31] X. Niu, B. M. Hemminger, C. Lown, S. Adams, C. Brown, A. Level, M. McLure, A. Powers, M. R. Tennant, and T. Cataldo. National study of information seeking behavior of academic researchers in the United States. *J. Assoc. Inform. Sci. & Tech.*, 61(5):869–890, 2010.

[32] S. Pontis and A. Blandford. Understanding "influence:" an exploratory study of academics' processes of knowledge construction through iterative and interactive information seeking. *J. Assoc. Inform. Sci. & Tech.*, 66(8):1576–1593, 2015.

[33] S. Pontis, A. Blandford, E. Greifeneder, H. Attalla, and D. Neal. Keeping up to date: An academic researcher's information journey. *J. Assoc. Inform. Sci. & Tech.*, 68(1):22–35, 2017. Online since 22 October 2015.

[34] F. Radlinski, M. Szummer, and N. Craswell. Inferring query intent from reformulations and clicks. In *WWW*, pages 1171–1172. ACM, 2010.

[35] S. Y. Rieh et al. Analysis of multiple query reformulations on the web: The interactive information retrieval context. *Inform. Proc. & Manag.*, 42(3):751–768, 2006.

[36] A. Shiri. Query reformulation strategies in an interdisciplinary digital library: The case of nanoscience and technology. In *ICDIM*, pages 200–206. IEEE, 2010.

[37] M. Shokouhi, R. Jones, U. Ozertem, K. Raghunathan, and F. Diaz. Mobile query reformulations. In *SIGIR*, pages 1011–1014. ACM, 2014.

[38] J. Tang. Aminer: Toward understanding big scholar data. In *WSDM*, page 467. ACM, 2016.

[39] J. Teevan. The re:search engine: Simultaneous support for finding and re-finding. In *UIST*, pages 23–32. ACM, 2007.