Investigating Query Reformulation Behavior of Search Users[⋆⋆⋆]

Jia Chen, Jiaxin Mao, Yiqun Liu, Min Zhang, Shaoping Ma

Department of Computer Science and Technology, Institute for Artificial Intelligence,
Beijing National Research Center for Information Science and Technology,
Tsinghua University, Beijing 100084, China

{chenjia0831,maojiaxin}@gmail.com,{yiqunliu,z-m,msp}@tsinghua.edu.cn

Abstract. Search engine users usually strive to reformulate their queries in the search process to gain useful information. It is hard for search engines to understand users' search intents and return appropriate results if they submit improper or ambiguous queries. Therefore, query reformulation is a bottleneck issue in the usability of search engines. Modern search engines normally provide users with some query suggestions for references. To help users to better learn their information needs, it is of vital importance to investigate users' reformulation behaviors thoroughly. In this paper, we conduct a detailed investigation of users' session-level reformulation behavior on a large-scale session dataset and discover some interesting findings that previous work may not notice before: 1) Intent ambiguity may be the direct cause of long sessions rather than the complexity of users' information needs; 2) Both the added and the deleted terms in a reformulation step can be influenced by the clicked results to a greater extent than the skipped ones; 3) Users' specification actions are more likely to be inspired by the result snippets or the landing pages, while the generalization behaviors are impacted largely by the result titles. We further discuss some concerns about the existing query suggestion task and give some suggestions on the potential research questions for future work. We hope that this work could provide assistance for the researchers who are interested in the relative domain.

Keywords: Query suggestion · Query Reformulation · User Behavior Analysis

1 Introduction

With the rapid development of Web search techniques, people are becoming increasingly dependent on search engines to solve problems these days. However, it seems users tend to submit short and ambiguous queries [15] which are too vague to be fully understood by search engines. This makes query formulation a bottleneck issue in the usability of search engines [10]. Sometimes the users may

⋆ This work is supported by Natural Science Foundation of China (Grant No. 61622208, 61532011, 61672311) and National Key Basic Research Program (2015CB358700).

⋆⋆ This research is partly supported by the Tsinghua-Sogou Tiangong Institute for Intelligent Computing.

endeavor several search rounds to reformulate their queries until they find some relevant results that fulfill their information needs. To ease the users' burden, modern commercial search engines usually provide query suggestions for them to better acknowledge their search tasks and express their queries more clearly. The previous work [16] has shown that appropriate query suggestions can significantly improve users' search satisfaction, especially for those navigational queries. Therefore, it is of vital importance to improve the query suggestion performance in commercial search engines.

To better understand users' search intents and model their information needs, session context information such as previous queries and click-through data have been employed by numerous approaches for query suggestion. Many of them make good use of the "wisdom of crowds" to mine the relations and similarities between queries. Some methods extract the query co-occurrences in the search log or apply Markov models to learn the query connections [2,17]. They assume that the most frequent follow-up queries for the current one are more likely to be submitted by users in the next. These methods are simple and achieve good performances sometimes, but may also suffer from the data sparsity problem (e.g. query suggestion for long-tailed queries). To tackle this obstacle, some researchers attempt to incorporate statistical features or clickthrough information into their models to learn better user reformulation behaviors. Jiang et al.[3] conduct a detailed analysis on a search log and extract some features to learn user reformulation behavior through some learning-to-rank algorithms. With the emergence of deep neural networks, more work focus on employing Recurrent Neural Networks (RNNs) to model users' sequential intra-session behaviors. For instance, Li at el.[8] propose a hierarchical attention network that applies the attention mechanism at both word- and session-level for context-aware query suggestion. Jiang at el.[7] first incorporate the embeddings learned from a session-flow graph into a reformulation inference network to predict the reformulation embedding for the next query.

Although these methods have achieved exciting performances in terms of predicting the next query within a search session, it is still unclear whether they can enhance users' search processes. In fact, to which extent they can help users to better, faster complete their search tasks still remains to be investigated so far. Moreover, existing evaluation approaches have their own limitations. Most studies aim to improve the rank of the next query in the candidate sets. However, one question is why should we boost the priority of this query even if it does not articulately express the user's search intent? Also, some other concerns should be taken into consideration. For example, there are a number of similar queries in the candidate list, but we only take one of them as our intended query and ignore the semantic similarities. Then the one-hit metrics such as MRR@k, SR@k or MISS@k may not accurately evaluate the suggestion performance under this circumstance. Therefore, we need to reacquaint the query suggestion task and search for the directions we should head towards.

To better investigate the context-aware query suggestion task itself, we conduct a meta-analysis on the session-level user reformulation behavior on a large-scale session data. The dataset totally contains more than 5,300,000 sessions extracted from a huge commercial search engine log [1]. In this paper, we propose three research questions as follows:

[1] To access the dataset, please contact chenjia0831@gmail.com.

– **RQ1**: When will people submit a long search session?
– **RQ2**: How does the user reformulation pattern evolve within a session?
– **RQ3**: How do users reformulate their next query when inspired by the previous search round? Can we find any relationship between some interaction signals (such as clicks) and users' reformulation behavior?

These questions are fundamental but crucial for exploring the query suggestion task. The analysis results may contribute to the redefining of the problem and the redesigning of better query suggestion metrics. We then present some of our considerations of the query suggestion task itself to provide references for other researchers who are interested in relative domains.

2 Related Work

There are a number of existing work aiming to optimize the ranks of the intended query (or the next query) in the candidate query set. Some work mine the inter-query dependencies through the co-occurrences in the query log, the query flow graph or the bipartite graph [10,19]. Recently, Sordoni at el.[5] first employ RNN for sequential query prediction and generation. To handle the data sparsity problem in utilizing the "wisdom of crowds", some researchers begin to exploit manually extracted features or clickthrough data for better user intent modeling. Wu at el.[6] propose a feedback memory network to take the clicks and skips in previous search rounds as positive and negative feedbacks respectively and incorporate them into the session-level embeddings. Jiang at el.[7] employ node2vec [20] to train the node embeddings from a session flow graph and then feed them into a reformulation inference network for query suggestion.

Instead of just boosting the system performance, some other studies make some efforts on understanding user reformulation behaviors. Liu et al.[18] analyze the nature of the query recommendation process from the user's perspective and use click-through rate and user click amount to evaluate the effectiveness of their proposed snippet click model. To gain precise and detailed insight into which terms the users show a particular interest in, Eickhoff at el. study query refinement using the eye-tracking technique [4]. Jiang at el. extract some heuristic features according to a user behavior analysis and apply the LambdaMart algorithm to learn user reformulation [3]. Except for analyzing user behavior, some Information Retrieval (IR) researchers focus on solving the query ambiguity problem. Shokouhi at el. not only utilize the context information within the session to provide unambiguous query suggestion but also propose a context-sensitive result fusion approach to improve the retrieval quality for ambiguous queries [14].

The difference between our work and the previous ones is: we make a detailed analysis of the session-level user reformulation behavior and the query suggestion task itself. We not only aim at exploring the reason why people submit long sessions, but also focus on the evolution of the user search pattern within the ongoing sessions. We further investigate to find the relationship between some signals and the user reformulation actions. Finally, we come up with some of our concerns for the query suggestion task and propose some potential directions for future work in this task.

3 Dataset

In this section, we will briefly introduce our dataset. We extract our session data from a log recorded by *Sogou.com*, which is a major Chinese commercial search engine. The log was sampled from April 1st to April 18th, 2015, containing abundant Web search data on the desktop. For each query, the URLs, vertical types and the click information (whether be clicked and the click timestamp) for all returned results and the user IDs are recorded. Similar to previous studies, we use the 30-minute gap as the session boundary to split the queries submitted by the same users into search sessions. We then discard those sessions with only one query or more than 10 queries, because there is no context information to be utilized in single-query sessions while sessions that are too long may contain much noise. Because our dataset is Chinese-centric, we adopt an open-sourced tool called `jieba` [2] for word segmentation. We randomly sample 10% of all the sessions as our testing set and the rest as the training set. To ensure the consistency, we abandon the testing sessions whose last query does not appear in the training set. Then we randomly sample 10% of the training data as our validating set for further system parameter tuning. Finally, there are 5,045,625 training sessions(including 505,036 validating sessions) and 331,605 testing sessions, respectively.

Table 1. Basic statistics of sessions with different context lengths.

Data\Context Length	**Short(1)**	**Medium(2-3)**	**Long(≥4)**	**All**
Training+Validating	3,261,183	1,393,824	390,618	5,045,625
Testing	223,467	85,324	22,814	331,605

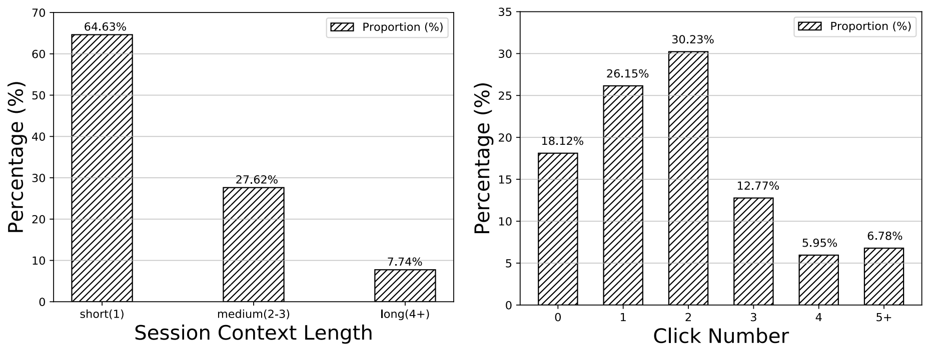
Table 1 presents the statistics of sessions with different context lengths in our dataset. To explore the system performance across sessions with different lengths, we split the session data into three groups, i.e. short sessions with only one previous query, medium sessions with 2-3 query contexts, and long sessions with more than four previous search rounds. From Figure 1(a), We can find that over 60% sessions are short sessions, which indicates that in the real-world Web search scenario people tend to submit only one query reformulation. We further explore the number of clicks within a session and present the statistical results in Figure 1(b). Over 80% of all sessions have at least one click. Generally, compared to another search engine log AOL [1], our dataset owns more short sessions but contains obviously more clickthrough information.

4 Session-level User Reformulation Behavior Analysis

In this section, we will make a detailed analysis of session-level user reformulation behavior on our training set (including the validating set). Here we leave the testing set as unknown. [3]

[2] https://pypi.org/project/jieba/

[3] Other researchers can use the testing set for system performance evaluation.



(a) Session length distribution (b) Session click distribution

Fig. 1. Distributions for session lengths and clicks.

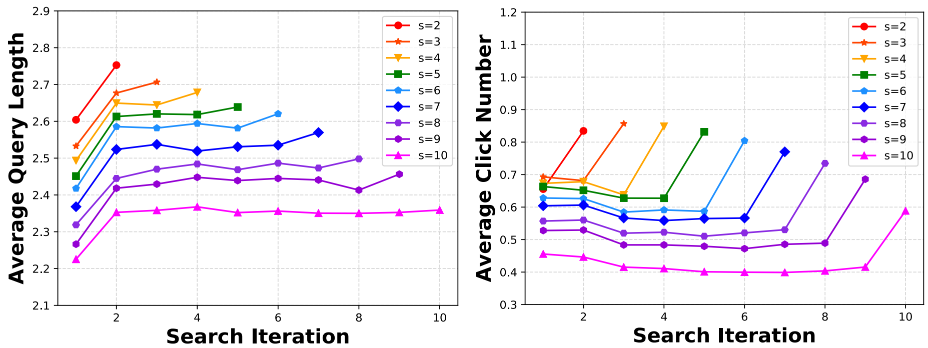
4.1 Analysis of Session-level User Reformulation Pattern

It is crucial to investigate users' session-level behavior to find heuristics for designing better query suggestion algorithms or evaluation metrics. In this subsection, we will study the following two research questions:

- **RQ1**: When will people submit a long search session?
- **RQ2**: How does the user reformulation pattern evolve within a session?

To answer the above questions, we will make a stratified analysis of the users' session-level search patterns on our session data.

A search session is consist of a sequence of queries $S = \{q_0, q_1, ...q_{|K|}\}$. The query lengths may change a lot at different session positions because of user intent shift. We present the trends of the query length and the user clicks across each search iteration in sessions with 2-10 queries in Figure 2(a) and 2(b), respectively. As shown in Figure 2(a), we find that generally the shorter a session is, the longer the queries within it are. This is different from the previous work [3], which reports longer sessions usually contain queries with more terms. We also observe that query lengths will always increase at the beginning, and then vibrate within a small range during the search processes. One possible explanation for this phenomenon is that although users may have complex information needs in longer sessions, they may not be able to express their needs clearly in the search query at the beginning and need to attempt multiple search rounds in a trial-and-error process to find an appropriate query expression. On the contrary, in short sessions, users may have higher-level cognition towards their information needs and thus can formulate their query with more terms and less ambiguity. We further explore the average click number in each session position. In Figure 2(b), we can observe a sharp rise of click number in the last two search rounds within a session in all lengths. There is an average click of over 0.6 in the last search iteration. This indicates that at the end of sessions, users click on the results and may be more satisfied with the search task so they choose to end the session. The slight decline of the curve in the middle of a session may be due to user intent shifts or expression reformulations. At this period, users tend to try



(a) Query length across session iterations (b) Click number across session iterations

Fig. 2. Trends of query length and click number across session iterations

different query expressions and will not click on the results until there are good results. From this point of view, long sessions may be mainly directly caused by the intent ambiguity rather than the complex information needs.

Query lengths and click numbers cannot provide details for user behaviors . Therefore, we compare the proportions of each reformulation type in short sessions and long sessions in our dataset, respectively. In the search process, users may learn from the search results and reformulate her next query. User reformulation behavior in our data can be normally divided into four main groups: *specification*, *generalization*, *repetition* and *others*. Specification includes those reformulations adding constraints to the original queries to narrow down the scope of search results. So the query lengths usually increase in this condition. On the other hand, users may also generalize the queries by loosening the search constraints and deleting some terms. Some other reformulation types include spelling change (or character change in Chinese), parallel shift, synonym, intent shift, and etc. Note that there are differences between the parallel shift and intent shift. In parallel shift, users may focus on various facets of the same object or problem. However, they may focus on two different objects when their search intents have shifted. In our session data, there are a proportion of query repetitions. We think that this is because the paging down operations by users are also recorded as new query submissions by the search engine. Some typical examples of each reformulation types in our data are presented in Table 2. Note that our session data is huge so we can only roughly label the reformulation type for every two continuous queries q_{t-1} and q_t according to the following definitions:

- **Specification** : $+\triangle q_t \neq \emptyset, -\triangle q_t = \emptyset$;
- **Generalization** : $+\triangle q_t = \emptyset, -\triangle q_t \neq \emptyset$;
- **Repetition** : $+\triangle q_t = \emptyset, -\triangle q_t = \emptyset$;
- **Others** : $+\triangle q_t \neq \emptyset, -\triangle q_t \neq \emptyset$.

where $+\triangle q_t = \{s | s \in S(q_t), s \notin S(q_{t-1})\}$, $-\triangle q_t = \{s | s \notin S(q_t), s \in S(q_{t-1})\}$, and $S(q)$ denotes the term set of a query q. Here we do not consider the semantic similarity, so the proportion of the "others" reformulation type we report can be higher than the ground truth value because a portion of specification or generalization cases might also be labeled as the "others" type.

Table 2. Typical examples of each reformulation types (translated from Chinese).

<table>
<tr><th colspan="2">Reformulation</th><th>Examples</th></tr>
<tr><td colspan="2">Specification</td><td>Minecraft → Minecraft skin websites</td></tr>
<tr><td colspan="2">Generalization</td><td>Transformers: Age of Extinction → Transformers</td></tr>
<tr><td colspan="2">Repeated Queries</td><td>Xiaomi→ taobao.com → taobao.com</td></tr>
<tr><td rowspan="4">Others</td><td>spelling change</td><td>Datong Securities(大同证券) → Datong Securities(大通证券)</td></tr>
<tr><td>parallel shift</td><td>Conan the movie version → Conan the mandarin version</td></tr>
<tr><td>synonym</td><td>Running Man 2(跑男 2) → Running Man 2(奔跑吧兄弟第二季)</td></tr>
<tr><td>intent shift</td><td>Ultra Magnus(通天晓) → The Fallen(堕落金刚)</td></tr>
</table>

Table 3 presents the comparison between short and long sessions in terms of the proportion of each reformulation type. We notice that there is no significant difference in the repetition action. This shows that there might be equal chances that users may repeat their last query or click the page-down button in short and long sessions. However, there is a rise of about 50% in the specification action from the long session condition to short session condition (i.e. from 9.93% to 14.14%). This gap is distributed to the generalization and the "others" reformulation types. We can learn from Table 2 that in generalization and the "others" reformulation types, users are more likely to shift their intents or expand the search scope. They choose to reformulate their queries in this way maybe because they have scanned the results of the current query and believe that this query is not heading for the desired information. Not sufficiently acknowledging their information needs and submitting queries with ambiguity to the search engines may be the direct cause of why people submit a long search session.

Table 3. Probabilities of each reformulation type in short and long sessions

Session Length	Generalization	Specification	Repetition	Others
Short(2)	0.0366	0.1414	0.3085	0.5136
Long(5+)	0.0480	0.0993	0.3042	0.5486

Having figured out the possible reason why users will submit a long session, we further explore how the user search pattern evolves within a session. The user search pattern is closely related with their reformulation behaviors, thus we calculate the proportions of each reformulation type across the session and plot the results in Figure 3. Generally, we notice there are much more repetition and other reformulation cases than either the generalization or specification at all steps. Due to its small proportion, it is hard to find any markable trends for users' generalization behavior across the session process. However, we find a stable decay (14% to only 7%) of the specification action from the first to the last reformulation step. This huge decline suggests that users are more and more clear about their search intents thus are less likely to add some constraints on their next queries to narrow down the search scope.

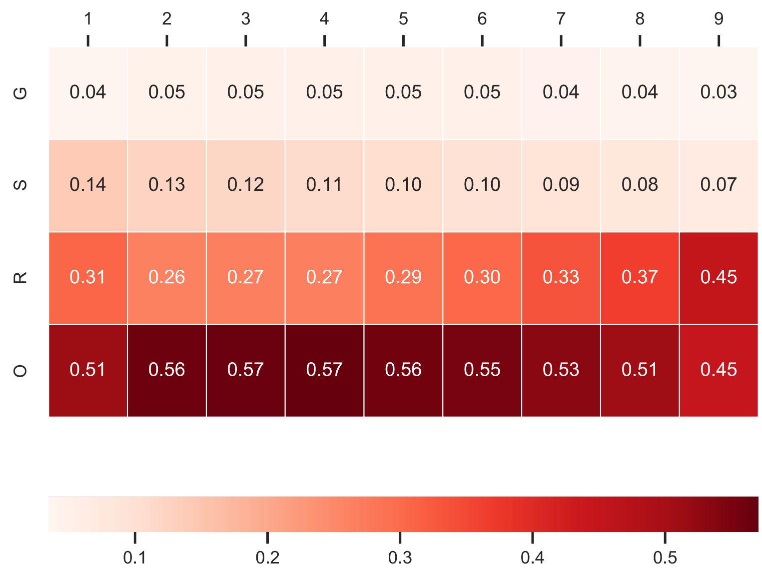


Fig. 3. The heatmap of the proportion of each reformulation pattern (G-generalization, S-specification, R-repetition, O-others) across reformulation steps within a search session. Here we take all sessions into consideration.

As for the "others" reformulation type, there is a sharp rise from the first to the second reformulation step, and then a slow decay from the fifth step to the eighth one, and finally a huge drop. This trend is exactly opposite to that of the repetition whose proportion first declines and then increases. The darker blocks in the fourth row of the heatmap indicate that people may shift their intents more frequently and submit more various queries in the middle of a session. In the last several search rounds, they tend to repeat their previous queries or examine more results. Finally, they end the search session because their information needs are mostly satisfied. From this analysis, we find that users' exploration process of searching for an appropriate query can be up to 7˜8 search rounds. Reflect on the current query suggestion task that mainly aims at optimizing the rank of the predicted query, a more urgent goal may be helping users to reformulate their queries with less ambiguity and shortening the search process.

4.2 User Reformulation Behavior

It is of vital importance to study the mechanism of how users reformulate their queries within a session. In this subsection, we will investigate the following research question:

– **RQ3**: How do users reformulate their next query when inspired by the previous search round? Can we find any relationship between some interaction signals (such as clicks) and users' reformulation behavior

We first compare the dwell time across each reformulation step in Figure 4(a). The results show that users spend more time browsing the results in shorter sessions. This finding is also consistent with our previous analysis that users tend to have a clear search intent in shorter sessions so they can receive better results and spend more time reading on them. In long sessions, the dwell time first declines and then increases, which implies users usually engage more at the beginning and the end of a session. There might be more attempts in the middle of a session.

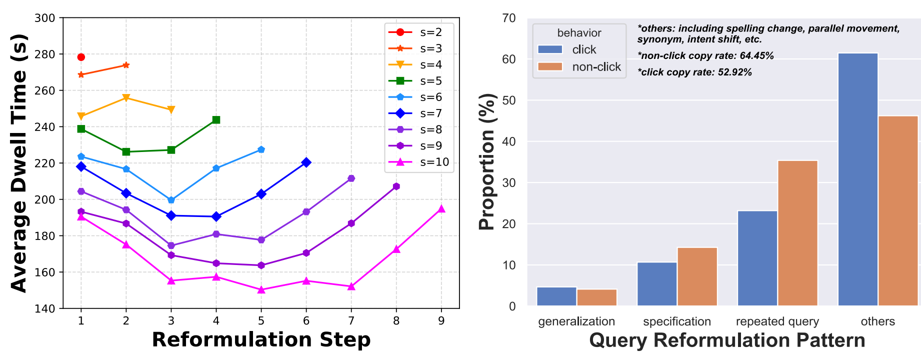
In the last subsection, we have shown that the search engine users will change their reformulation patterns during the search process. Inspired by some contents in the previous search round, users will rewrite the following query to obtain more appropriate results. To make sense how users can be influenced by the previous search round and reformulate their next queries, we crawl the titles for each search result in our data and analyze the relationship between these titles and the newly-formed query. For over 92% queries, we have crawled the titles for at least five results. We then count up the number of the cases in which the terms added in the specification action or deleted in the generalization action also appear in the result titles of the previous query. All cases are divided into four groups according to whether the result has been clicked or not: *click*, *skip*, *non-click*, and *others*. Here we define the skips as those results that have not been clicked but ranked higher than the last clicked result. Other results without clicks are denoted as the non-clicks. Users may also be impacted by other contents such as the texts in the search snippets, the landing pages, and etc. We regard these conditions as others.

The statistics are shown in Table 4. Note that there are less long sessions than short sessions, so the numbers of cases always decay across the reformulation steps. An interesting finding is: both $+\triangle q_t$ and $-\triangle q_t$ are more likely to be influenced by the clicked results than the skipped results. To our surprise, the probability of a user being affected by the clicked results is around five times of that by the skipped ones if she deletes some words from her previous query. However, this margin drops to only three times in the case of the added terms in the specification condition. This finding indicates that clicks are not always the positive feedback signals for user information needs. Users may first check the results, realize the current query is not suitable for their search purposes, and then delete some terms from the current query according to the result titles. Another finding is that the users' reformulation behavior can also be largely impacted by those non-click titles. Especially, the non-click cases account for more than a half of the amount in the generalization condition. One possible reason may be: although the users do not click these results, they are likely to examine them and somehow judge them as not relevant. In addition, the "others" condition occupies a much larger proportion in term of $+\triangle q_t$ than $-\triangle q_t$. This gap suggests that the user specification behavior can be influenced more by other elements such as the snippets, the landing page contents, and etc. In contrast, they may not engage too much in the generalization case. They may just scan the result titles and then decide to delete some terms if they estimate that the current query is not appropriate or too specified.

Table 4. Statistics for the cases in which the terms in $+\triangle q_t$ and $-\triangle q_t$ also appear in the result titles of the previous query. Here we only consider the results whose titles we have successfully crawled.

Condition		Reformulation Step								
		1	2	3	4	5	6	7	8	9
$+\triangle q_t$	*clicks*	11,890	4,223	1,655	756	401	185	76	40	23
	skips	3,494	1,479	596	255	130	59	29	20	9
	non-clicks	39,669	14,120	6,058	2,897	1,485	759	365	172	83
	others	109,670	39,834	17,764	8,706	4,552	2,439	1,296	659	271
$-\triangle q_t$	*clicks*	48,646	18,664	7,634	3,675	1,881	922	473	250	104
	skips	9,822	4,450	1,940	901	464	234	122	62	28
	non-clicks	98,200	36,939	16,131	7,923	4,120	2,172	1,076	592	233
	others	27,998	10,829	4,889	2,441	1,244	678	404	174	71

To further investigate the relationship between clicks and user reformulation behavior, we calculate the proportion of each kind of reformulation type given the previous query causes clicks or not. The results are shown in Figure 4(b). We can learn that for the queries do not cause clicks, there is a higher probability(64.45%) that the user may copy some terms from these queries to reformulate her next query. However, the probability drops to only 52.92% if there is no click in the previous search round. This find is consistent with the results reported in previous work [3]. On one hand, users are more likely to reformulate specified or repeated queries if she clicks no result of the current query. One possible explanation is that users do not find any relevant results so they tend to click the page-down button or add some constraints on the current query to search for more relevant results. On the other hand, there is a higher probability of other reformulation behaviors when people click some results. In this condition, people may find relevant documents and feel satisfied to some extent. Therefore, they may tend to end up the current search subtask or even shift their intents to start another subtask.



(a) Dwell time across reformulation steps (b) Distribution of reformulation patterns

Fig. 4. Some statistics of users' session-level reformulation behaviors.

4.3 Conclusions

To sum up, the above analysis suggests that:

- It may be the intent ambiguity or expression difficulty that directly cause longer sessions rather than users' complex information needs.
- Users tend to specify their queries in the first reformulation step to narrow down the search scope, and then continue their search processes in a trial-and-error manner.
- Surprisingly, for both the added terms $+\triangle q_t$ and the deleted terms $-\triangle q_t$ in the reformulation actions, users are possibly impacted more by the clicked results than the skipped ones.
- Users can be influenced more by the result titles when they decide to delete some terms in the current query, while they may be more likely to refer to other contents such as the result snippets or the texts in the landing pages for query specification.
- If a user clicks some results in this search round, then there is a higher probability she will copy some terms from the current query to reformulate her next query.
- Users will click more documents at the end of a session. They may find some relevant documents, feel satisfied with them, and then end their search processes.

From the above analysis, we find several assumptions in some previous work may not be so meticulous and should be further improved.

5 Discussions and Future Work

In this paper, we make a detailed investigation of users' session-level reformulation behavior. We find that some of the assumptions adopted by the previous work are not accurate enough and may hurt the robustness of their theory. Some main concerns for the query suggestion task are:

Firstly, since there are a proportion of long sessions, many users may endeavor several search rounds until they find an appropriate query. Since the query ambiguity and the expression difficulty will cause users struggled in a long session, a possible future work may be to shorten the search process by disambiguation rather than just to predict the next query. Also, the existing evaluation metrics such as MRR and SR have their limitations as they do not take the semantic similarities into consideration. More robust evaluation system should be constructed so far.

Secondly, according to our analysis results, titles of the clicked results can have great impacts on both users' specification and generalization actions. There may be problems if we just regard the click signals as the positive feedbacks and the skips as the negative feedbacks. More evidences should be collected for designing better query suggestion algorithms.

Last but not least, we have found that there are obvious marginal effects on the user clickthrough and reformulation behaviors. So the session boundary detection can also be an issue. Most existing work roughly adopt the 30 minutes as the threshold to split the query sequence into sessions and then continue their experiments on these sessions. However, when it comes to practice, it is hard to know when the user will end their search processes. Therefore, it is also crucial to predict the end of a session.

References

1. Brenes D J, Gayo-Avello D. Stratified analysis of AOL query log[J]. Information Sciences, 2009, 179(12): 1844-1858.
2. He Q, Jiang D, Liao Z, et al. Web Query Recommendation via Sequential Query Prediction[C]. international conference on data engineering, 2009: 1443-1454.
3. Jiang J Y, Ke Y Y, Chien P Y, et al. Learning user reformulation behavior for query auto-completion[C]//Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval. ACM, 2014: 445-454.
4. Eickhoff C, Dungs S, Tran V, et al. An Eye-Tracking Study of Query Reformulation[C]. international acm sigir conference on research and development in information retrieval, 2015: 13-22.
5. Sordoni A, Bengio Y, Vahabi H, et al. A Hierarchical Recurrent Encoder-Decoder for Generative Context-Aware Query Suggestion[J]. conference on information and knowledge management, 2015: 553-562.
6. Wu B, Xiong C, Sun M, et al. Query Suggestion With Feedback Memory Network[J]. the web conference, 2018: 1563-1571.
7. Jiang J, Wang W. RIN: Reformulation Inference Network for Context-Aware Query Suggestion[C]. conference on information and knowledge management, 2018: 197-206.
8. Li X, Liu Y, Li X, et al. Hierarchical Attention Network for Context-Aware Query Suggestion[C]. asia information retrieval symposium, 2018: 173-186.
9. Dehghani M, Rothe S, Alfonseca E, et al. Learning to Attend, Copy, and Generate for Session-Based Query Suggestion[J]. conference on information and knowledge management, 2017: 1747-1756.
10. Cao H, Jiang D, Pei J, et al. Context-aware query suggestion by mining click-through and session data[C]. knowledge discovery and data mining, 2008: 875-883.
11. Baryossef Z, Kraus N. Context-sensitive query auto-completion[J]. the web conference, 2011: 107-116.
12. Chen W, Cai F, Chen H, et al. Personalized Query Suggestion Diversification[C]. international acm sigir conference on research and development in information retrieval, 2017: 817-820.
13. Anagnostopoulos I, Razis G, Mylonas P, et al. Semantic query suggestion using Twitter Entities[J]. Neurocomputing, 2015: 137-150.
14. Shokouhi M, Sloan M, Bennett P N, et al. Query Suggestion and Data Fusion in Contextual Disambiguation[J]. the web conference, 2015: 971-980.
15. Gao B J, Anastasiu D C, Jiang X, et al. Utilizing User-input Contextual Terms for Query Disambiguation[C]. international conference on computational linguistics, 2010: 329-337.
16. Song Y, Zhou D, He L, et al. Post-ranking query suggestion by diversifying search results[C]. international acm sigir conference on research and development in information retrieval, 2011: 815-824.
17. Dang V, Croft B. Query reformulation using anchor text[C]. web search and data mining, 2010: 41-50.
18. Liu Y, Miao J, Zhang M, et al. How do users describe their information need: Query recommendation based on snippet click model[J]. Expert Systems With Applications, 2011, 38(11): 13847-13856.
19. Huang C K, Chien L F, Oyang Y J. Relevant term suggestion in interactive web search based on contextual information in query session logs[J]. Journal of the American Society for Information Science and Technology, 2003, 54(7): 638-649.
20. Grover A, Leskovec J. node2vec: Scalable feature learning for networks[C]//Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2016: 855-864.