Long Query Suggestion List: Prioritized or Organized?

Shuo Wang[1], Feng Jing[1], Jiangming Yang[1], Jibo He[2],

[1] Microsoft Research Asia, Beijing Sigma Center, Beijing, P. R. China
{shuowang, fengjing, t-lancey}@microsoft.com
[2] Peking University, Beijing, P. R. China, hejibo@gmail.com

Abstract. Query suggestion, offering related queries to users' original queries, is proven as an effective solution for users' difficulty in query formulation. However, the benefits of query suggestion are limited by difficulties in presentation of long query suggestion list. To address this issue, we proposed two approaches: (1) Prioritized, split the most searched and the rest to up and bottom respectively; (2) Organized, categorize the terms as refined and related, and cluster the terms into those two categories by similarity. User study shows organized presentation of long query suggestion list is more efficient and preferable than Prioritized presentation.

Keywords: Query suggestion, long list, search result clustering, user study

1   Introduction

Query suggestion (QS) is a proven solution for short, general, and ambiguous queries. In particular, it helps users who are poor in query formulation, and provides a larger scope of queries to choose from [4].

Despite of the benefits of query suggestion, many existing obstacles still remain unsolved, spoiling both the quality and user experience, i.e., the organization of long list query suggestion, and the tradeoff between number of query suggestions and cognitive load [6]. Different presentations of query suggestions affect the search tasks, and a more structured organization improves the efficiency of query suggestions [2].

AltaVista limits the number of the suggested terms to twelve both to conserve space on the page and to minimize cognitive load [5]. Yahoo [10] folds the suggestions into three levels with both "more" and "show all" links (Fig 1 Left). Ask [8] not only folds the suggestions, but also classifies them into three categories (Fig 1 Right).

The observation of eighteen users using Ask and Yahoo query suggestions revealed the following problems regarding to the query suggestion user interface:

- **Folding at the cost of coverage**

Many potentially relevant terms cannot be displayed in Yahoo; reducing the chance of an intersection with the users' specific information needs [5].

- **Clicking not necessarily worth the effort**

Yahoo usually only offers one to three more terms for every click on "more" and "show all". On the contrary, Ask often surprises users with an overwhelming long list (as many as 31 more suggestions for "tom cruise" in the "*Related Names*"). Too few or too many queries both can do limited help for users.

- **Ignoring as a result of misunderstanding**

The query suggestion of Ask, presented in a vertical list on the right, could be misunderstood as advertisement. Though right above the results, the query suggestion of Yahoo is too obscure to draw attention for its visual similarity and crowdedness.

Fig. 1. Yahoo: clicking twice to unfold the list (Left). Ask using a "more" link to hide the long list[1] (Right).

The hierarchical structure [3], fixed number list [5], and graphical visualization [9] are all existing methods of query suggestion presentation, but none of them have an in-depth investigation of the long list issue.

2   Prioritized vs. Organized

We proposed two prototypes as possible solutions for the unorganized long list of query suggestions: *Prioritized* and *Organized*.

- **Prioritized** (Fig 2):

The *prioritized* approach relies on the assumption that users generally share a large portion of search needs. Popular queries and related query suggestions cover a significant part of the web searches.

All the query suggestions are categorized into "*most searched*" and "*related searches*" by popularity. The relatively small numbers of queries (less than 100 characters per line), labeled as "*most searched*" at the top of the search results, satisfy most of the search needs without significant increase in browsing and cognitive load. The suggestion in the "*related searches*" at the bottom, serves as complementary queries to ensure the coverage and relevance of query suggestions.

- **Organized** (Fig 3):

According to the cognitive processing theories, cognitive chunks, that is, a semantically or functionally similar suggestion cluster exert no significantly load than

---

[1] The three categories were originally arranged in vertical on the right side of the page.

a single query [7]. Therefore, two measures are applied to display the query suggestion more *organized* or in cognitive chunks:

1. Query suggestions are classified by function of "refine" or "expand" the search result respectively into "*Refine by*" and "*Also try*". "*Refine by*" consists queries formed by the original query and a modifier and possible refinements; "*Also try*" includes related queries, containing no or only part of the original query.

2. Queries are organized in clusters by semantic similarity with the symbol of "|" to separate two neighboring clusters.

Fig.2. Prioritized: "Most searched" and "Related searches" (the results are trimmed for space consideration).

Fig. 3. Organized: Query suggestions organized as "Refine by" and "Also try".

3   User Test

A 3 X 3 within-group design was applied in this user test to compare the usability of the three prototypes of search engines (no query suggestion, *prioritized* and *organized* presentation of query suggestion) in three types of search scenarios, including informational, navigational and transactional, according to the search taxonomy [1].

All 18 participants had at least 6-month experience of web search engine, and all had used the feature of query suggestion in existing search engines. There were nine tasks in total (three minutes each): three tasks for each search scenario.

Informational: gather information on a subject. Participants are required to find answers to five questions regard to "Anemia", "Dyslexia", and "Chalazion";

Navigational: find specific web sites. The tasks of navigational scenario in the study are searching for websites of three given celebrities;

Transactional: searching for the download sites of "windows vista", "firefox", and "windows live messenger".

Latin Square design and counterbalance of both search engine and task type sequences were applied to minimize possible learning effect and sequence effect.

4   Result and Discussion

A 3 X 3 ANOVA tests were applied to compare user performance with the three prototypes under three search scenarios. As in Fig 4, both search engine type ($F_{(2,153)}$ = 28.544, $p<0.05$) and task type ($F_{(2,153)}$ = 5.861, $p<0.05$) had significant main effect on performance. Moreover, the influence of search engine type on performance was independent of the task type for the insignificant interaction effect ($F_{(4,153)}$ = 1.721, $p>0.05$).

Post-Hoc analysis of the main effect of search engine type showed (Fig. 4) that both Prototype *Organized* (M=4.537, SD=0.335) and *Prioritized* (M=4.852, SD=0.335) were significantly effective than the prototype of Non-suggestion (M=3.315, SD=0.335). This result once more validated the necessity of query suggestion in web search and supported both the feasible and usefulness of our algorithms and interface designs.
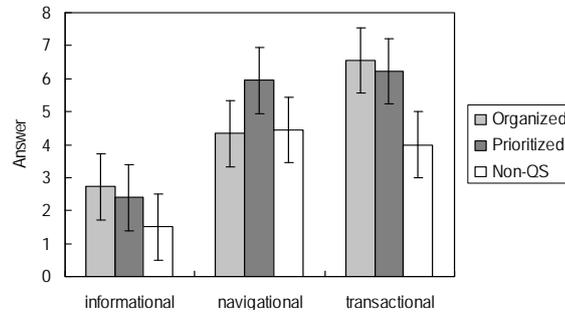

Fig. 4. Comparisons of performance with three prototypes.

There was no statistically significant difference between *Organized* and *Prioritized*, but there was a tendency that *Organized* had slight advantage in the task of "informational" and "transactional", while *Prioritized* was more suitable for "navigational" task.

Questionnaires and Paired-sample t-tests were applied to compare user experience: "usefulness", "comprehensibility", "easy of use", "number of suggested terms" and "overall satisfaction". Table 1 showed that *Organized* excelled over *Prioritized* in all.

Table 1. Comparison of User Experience between *Prioritized* vs. *Organized*.

User Experience	Difference Mean	SD	t $_{(17)}$	*p*
usefulness	0.611	1.145	2.265	0.037*
comprehensibility	0.833	1.689	2.093	0.052**
easy of use	0.500	0.857	2.474	0.024*
numbers	0.222	0.943	1.000	0.331
satisfaction	0.444	0.856	2.204	0.042*

(Note: * significant at .05 level; ** significant at .10 level)

5   Conclusions

User study shows that the *Organized* prototype, with both classification of query function and query clustering, is efficient for three kinds of search needs, especially in "informational" and "transactional" search tasks. The *Organized* prototype also offers more satisfactory user experience than the *Prioritized* prototype.

References

1. Andrei Broder, A taxonomy of web search, *ACM SIGIR*, v.36 n.2, Fall 2002

2. E. N. Efthimiadis. Interactive query expansion: A user-based evaluation in a relevance feedback environment. *Journal of the American Society for Information Science*, 51(11):989–1003, 2000.

3. Hideo Joho, Claire Coverson, et al, Hierarchical presentation of expansion terms. *SAC* (2002), 645-649

4. Jansen, B.J., et al. Real life, real users and real needs: A study and analysis of users' queries on the Web. *Information Processing and Management*, 36(2), 207-227.

5. Peter Anick, Using terminological feedback for web search refinement: a log-based study, In *Proc. SIGIR*, ACM Press, July 28-August 01, 2003

6. Peter Bruza, et al, Interactive Internet search: Keyword, directory and query reformulation mechanisms compared. *In Proc of ACM, SIGIR 2000*, 280-287.

7. Simon, H. A., How big is a chunk? *Science* 183, 482-488

8. Ask query suggestions. http://www.ask.com

9. Quintura. http://www.quintura.com/

10. Yahoo Query suggestions. http://search.yahoo.com