The perfect search engine is not enough: a study of orienteering behavior in directed search

Conference Paper · January 2004

DOI: 10.1145/985692.985745

8 authors, including:

David Ron Karger
Massachusetts Institute of Technology
388 PUBLICATIONS **42,895** CITATIONS

SEE PROFILE

The Perfect Search Engine is Not Enough: A Study of Orienteering Behavior in Directed Search

Mark S. Ackerman, Christine Alvarado, David Karger & Jaime Teevan

What: Sometimes, searching for electronic information can be a complex, multistage process, where a user's information need evolves throughout the course of the search. On the other hand, often a user's search target is known in advance (e.g., a phone number or address). To understand how people perform these small, directed searches, we performed a modified diary study that investigated how people performed personally motivated searches in their email, in their files, and on the Web. Although earlier studies of directed search focused on keyword search, most of the search behavior we observed did not involve keyword search. Instead of jumping directly to their information target using keywords, our participants navigated to their target with small, local steps using their contextual knowledge as a guide, even when they knew exactly what they were looking for in advance. This stepping behavior was especially common for participants with unstructured information organization. The observed advantages of searching by taking small steps include that it allowed users to specify less of their information need and provided a context in which to understand their results.

Why: Researchers have tried to support directed search by attempting to build a "perfect" search engine–i.e., one that returns exactly what is sought given a fully specified information need. Attempts to build such a search engine have focused on improving on keyword search by permitting users to better specify their information need through meta-data [6], natural language [3], and even context [2]. Such a perfect search engine, although perhaps impossible to flawlessly construct, could solve many problems people currently have with search tools.

However, even when a person knows exactly what they are looking for, the perfect search engine might not be enough. Consider Rachel, a participant in our study. She attempted to locate a document that she knew existed in her file system. Although she knew exactly what document she was looking for (i.e., her information need was not evolving), she could not describe the document, its contents, or its location in advance:

```
I don't know how I could have the directory [the document was in] in mind without
knowing its name, but I felt sure which it was.
```

Because she could not specify her information need, a "perfect" search engine probably would not have helped her. Nonetheless, she successfully found her target through a series of small steps, using the local context at each stage of her search to inform her next step.

How: We performed a qualitative observational study examining what people did when working with their email, their file system, and the World Wide Web. We did not set about to test specific hypotheses, but rather to understand our participants' behavior. We wanted to understand what people did with their electronic information, and to do this we conducted a series of semi-structured interviews in which participants reported their information activities twice a day over the course of a week. The interviewer interrupted participants' work at unspecified times and prompted them to describe recent activity with their emails, their files, and on the Web. This method enabled us to gain a deep understanding of naturalistic search behavior because participants worked with their own information. Our method was similar to the diary studies used in many information seeking studies, as well as the Experimental Sampling Method [4].

During each semi-structured interview we asked the participants to describe what they most recently "looked at" and what they most recently "looked for" in their email, files and Web. What precisely defined "looking for" versus merely "looking at" was defined by the participants themselves based on what they considered effort. Overall, we obtainted 151 interviews. In addition, we conducted longer semi-structured interviews (1 hour) with each participant about their information and conducted some direct observations. The data were analysed using standard qualitative techniques (e.g., [1]).

Progress: We observed that when people searched for specific pieces of information, such as phone numbers or addresses, they generally knew exactly what they were looking for at the onset of their

search. We expected our participants to take advantage of this advanced knowledge of their target by using keyword search more often than they would when searching for general information, where the information need often evolves. Surprisingly, only 34 of the 81 searches for specific information that we observed (42%) involved keyword search, compared to 23 of the 42 searches for general information (55%). To understand how our participants performed directed searches, and why they avoided keyword search in many cases, we performed a qualitative examination of our data and uncovered two differing search strategies: *orienteering* and *teleporting*.

We observed many directed searches, like the following, where a series of small steps were used to narrow in on the target. Here, although Jim is looking for the office number of a math professor, Connie Monroe, he does not try to find it directly but instead looks for it via her department's page.

Interviewer: Have you looked for anything on the Web today?
Jim: I had to look for the office number of the Harvard professor.
I: So how did you go about doing that?
J: I went to the home page of Math Department at Harvard.

This participant then goes on to explain that he knows there is a specific Web page with the address:

I: Did you know it would be there [on a page] or you just hoped it would be there?
J: I knew that she had a very small Web page saying, I'm here at Harvard.
Here's my contact information.
I: So you went to the Math department, and then what did you do over there?
J: It had a place where you can find people, a link to the page where you can find people and I went to that page and they had a dropdown list of visiting faculty, and so I went to that link and I looked for her name and there it was.

This search by localized or situated navigation is an illustration of what we call *orienteering*. Orienteering involves using both prior and contextual information to narrow in on the actual information target, often in a series of steps, without specifying the entire information need up front. We observed that orienteering was heavily relied upon, even in directed search for specific information.

At the other end of the spectrum from a search strategy that involves many local, situated steps is a strategy we call *teleporting*. When a person attempts to teleport, they try to jump directly to their information target. Teleporting represents the behavior many search engines try to support in their quest to be "perfect". For example, if Jim, instead of browsing to Monroe's office number, had performed a search for, "Connie Monroe, office number," the perfect search engine would have brought him her office number.

Attempts at teleporting were surprisingly rare, and in our analysis, we investigated why people often chose not to teleport. In our data we found that three properties of orienteering offered our participants advantages over teleporting: it decreased their cognitive load, allowed them to maintain a sense of location during their search, and gave them a better understanding of their search result.

Future: We plan to further examine the nature of the contextual information used when orienteering. What people used to search for information appeared to vary based on whether they had seen the information before or not, as well as what type (e.g., email, file, Web) of information being searched for. We will use what we learned to inform the design and development of a next generation information management system [5]. As the amount of information we interact with grows, information management will increasingly become a problem we must deal with. Our study revealed behavioral patterns we can examine further in order to build tools to make this interaction more manageable in the future.

Research Support: This research is supported by NTT, HP, the Packard Foundation, the Oxygen Partnership and the National Science Foundation.

References:

[1] M. S. Ackerman and C. Halverson. Considering an organization's memory. In *Proceedings of CSCW '98*, 1998.

[2] S. Lawrence. Context in Web search. *IEEE Data Engineering Bulletin*, 23(3), 2000.

[3] J. Lin, D. Quan, V. Sinha, K. Bakshi, D. Huynh, B. Katz, and D. R. Karger. The role of context in question answering systems. In *Proceedings of CHI '03*, 2003.

[4] L. Palen and M. Salzman. Voice-mail diary studies for naturalistic data capture under mobile conditions. In *Proceedings of CSCW '02*, 2002.

[5] D. Quan and D. R. Karger. Haystack: A personalized, unified interface to creating, browsing, and organizing arbitrary information. In *Proceedings of CHI '04 (Extended Abstract)*, 2004.

[6] K. P. Yae, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *Proceedings of CHI '03*, 2003.