Paper Critique Ryen W. White (ryen@umd.edu), 4 February 2005

Wang, P. and Soergel, D. (1998). A Cognitive Model of Document Use during a Research Project. Study 1. Document Selection. *Journal of the American Society for Information Science*, 49(2), 115-133.

Overview

1. *What's the paper about?*
The paper presents the findings and analysis of part one of a longitudinal study into how users select and use documents. It describes the selection of documents as a decision process whereby users must make choices on the potential value/worth of documents based on the surrogate information provided by Information Retrieval (IR) systems. The study does not simply analyse final document selection decisions, but the processes that lead to the decision. It uses data gathered on the process and structure of document selection during a real life situations to establish what components affect user's decisions and the cognitive processes that occur during the selection. The data were gathered based on the verbal reporting of experimental subjects before, during (using a think aloud protocol) and after the project. The authors use this data to construct a cognitive model of decision-making in the selection process, based on information provided by surrogates and the assessment of document value that leads to a decision about how the document could be used. The authors claim that the model accounts for the use of personal knowledge and decision strategies applied in the selection process. They present a set of implications on how the model can be applied to the development of an intelligent system for helping users select documents.

2. *What are the contributions?*
The paper presents a model of the document selection process that may be useful to better understand how users search and has implications for the design of agents to improve the effectiveness of document retrieval. The main purpose of the paper was to use data from those with expertise and experience in document selection to improve the mechanisms used by IR systems to select documents and to improve how retrieval results are presented by default and facilitate the customization of this activity. Other contributions aside from the cognitive model of the document decision process include:

- A deeper understanding of how users select documents for later use
- A broad taxonomy of relevance criteria
    - where topicality was not always the basis for positive judgments
    - personal knowledge and document value are separate components of the model, not part of criteria
- A more complete understanding of:
    - what document representations were important
    - the extent to which each representation was important
    - what representations that were not available are potentially important

3. *What is the main line of argument in the paper?*
The paper argues that better document retrieval systems can be developed through the construction of a cognitive model of the document selection process. The authors assert that the judgment of relevance is not the same as the decision to use a document; for real users, the real task is to make a decision, not only a judgment of relevance. They therefore use human decision making and satisficing behaviour as the basis for their cognitive model of document selection through expanding the lens model and notions of information scent. The process involves four parts starting from when documents are represented by document information objects (DIEs) that are used to judge the document based on criteria (from pilot analyses of verbal reporting), which are applied to judge document values (from consumer theory) that lead to a selection decision. The process presented by the authors is governed by decision rules that determine the volume of information processed before a decision is reached. The authors argue that they can capture the processes behind this selection through a detailed analysis of verbal reporting to solicit the natural thoughts of each individual.

4. *What is the paradigm?*
The paper breaks slightly from the traditional paradigm of studies into information seeking behaviour and information use by focusing on the document selection process, rather than relevance judgment. Traditionally, studies of relevance research have focused on determining a set of relevance criteria across situations and users; the focus of the authors on decision making and a cognitive perspective on relevance fosters the development of new techniques for studying relevance judgments that take the process beyond the judgments themselves.

Significance and originality

1. *Were the ideas novel or original at the time?*
The analysis of relevance from a user-oriented and cognitive perspective was not original at the time. However, the strict separation of document information elements and relevance criteria used by users was original, as was:

- the richer set of relevance criteria that emerged as a result of the study
- the larger number of document representations that were used in the study
- the cognitive model of the document selection process
- the implications of the model for the design of IR system algorithms and interfaces

2. *Are the contributions significant?*
This work makes a significant contribution to our understanding of what criteria, values and document representations are important to users in the selection of documents in IR systems.

3. *What are the advantages and limitations of the approach?*
Advantages

- Insight into the thoughts of users as they make document selections
- Improved understanding of the relative importance of criteria, values and document representations
- Large user sample and diverse information needs

Limitations

- Qualitative data were unsystematic and incomplete
- Realism of "anchored judgement" (i.e., user gets to check all selections twice)
- Findings are not generalisable and the model needs to be further tested
- Low number of documents retrieved per user (i.e., average 52)
- Results are indicative not conclusive

Soundness

1. *Are the ideas that are presented technically correct?*
The ideas that are presented appear technically correct and the methodology is well structured and thought out.

Empirical evaluation

1. *Are the ideas evaluated using a study design that is appropriate for questions that the authors wished to answer?*
Since the research questions posed by the authors require insight into the cognitive processes of the user it seems reasonable to use a procedure such as think aloud to do so.

2. *If some assumptions are made, are they realistic?*
The authors assume that they will be able to better understand cognitive processes through think alouds and that users will be able justify their selection decisions in a way that is analyzable. The also assume that they can make inferences about the meaning of a statement from it's textual content and do not make reference to the use of audio information to resolve ambiguous meaning.

Related work

1. *Is the paper properly situated with respect to related work?*
Yes, the paper describes related work and situates itself in the context well.

Readability

1. *Is the paper well structured?*
The paper is well structured and is presented in a coherent way.

2. *Is the paper well written?*
The paper is well written, although the reader may find the 'Results and Findings' section somewhat overwhelming and may skip the details. The use of language is good and the ideas are presented in robust and comprehensive way.

Some things for you to think about…

This is my take on the paper, yours may be completely different. You may also want to consider issues such as: If you were reviewing this paper for publication, what would you recommend? Given the arguments and findings, would you have written the paper differently? How does this all relate to the Web? For example, do the findings on DIEs mean something for how Web search results should be presented? How about intelligent Web search agents?

Appendix

You may also be interested in some comments from the first author, Peiling Wang, who provided these insights into the good a bad aspects of the work from her perspective (thought this may be interesting for you):

Good points

The paper did well in distinguishing between the "criteria" users used to make decisions and the "document information elements" that provided the basis for users' decisions.

Many of the published relevance studies have mixed document information elements with user criteria. For example, Author was a user criterion by some other studies. My study showed that author provides information to infer quality, topic, etc. Therefore, quality and topic are criteria, but author is information element (attribute) of the document.

Bad points

I wish I had the word "relevance judgments" in title: A Cognitive Model of Document Use during a Research Project: Relevance Judgments and Document Selection