CHAPTER 9

IMPLICIT FEEDBACK: USING BEHAVIOR TO INFER RELEVANCE

DIANE KELLY
University of North Carolina at Chapel Hill

1. INTRODUCTION

The concept of relevance has a rich history in information retrieval (IR) that dates back well over 40 years (Borlund, 2003) and is necessarily a part of any theory of information-seeking and retrieval. Relevance feedback also has a long history in IR (Salton, 1971) and is considered an important part of interactive IR (Spink and Losee, 1996). Relevance feedback techniques often require users to explicitly provide feedback to the system, by, for instance, specifying keywords, selecting, evaluating and marking documents, or answering questions about their interests. The feedback that users provide during these interactions has been used for a variety of IR techniques and applications including query expansion, term disambiguation, user profiling, filtering and personalization.

Empirical studies have led to the general finding that users of interactive IR systems desire explicit relevance feedback features and, in particular, term suggestion features (Beaulieu, 1997; Belkin et al., 2001; Koenemann and Belkin, 1996). However, much of the evidence from laboratory studies has indicated that relevance feedback features are not used. While users often report a desire for relevance feedback and term suggestion, they do not actually use these features during their searching activities. Several reasons can be given for why this disparately exists. Users may not have additional cognitive resources available to operate the relevance feedback feature. While the extra effort required to operate the feature may seem trivial, the user is already potentially involved in a complex and cognitively burdensome task. Increased effort would be required for both learning the new system and operating its features. When features require more effort and additional cognitive processing than they appear to be worth, they may be abandoned all together. Furthermore, if relevance feedback features are not implemented as part of the routine search activity, they may be forgotten, no matter how helpful they are. This research, in part, has lead to the general belief that users are unwilling to engage in explicit relevance feedback.

Recently (Anick, 2003) demonstrated in a web-based study, that users made use of a term suggestion feature to expand and refine their queries, thus things may be changing. These results suggest the potential of term suggestion features in some types of information-seeking environments, especially for single session interactions. Hence it may just be the case that traditional relevance feedback interfaces have not effectively elicited feedback from users or optimally integrated relevance feedback features into current information interaction models.

Amanda Spink and Charles Cole (eds.), *New Directions in Cognitive Information Retrieval.* 169–186.

However, there are some problems with the sustainability of using explicit relevance feedback to personalize information-seeking interactions and use over time. These techniques typically require users to specify keywords or spend time selecting and marking documents. When considered from a long-term perspective, it seems unrealistic to expect users to read and rate a large number of documents on a daily basis, if they are even willing to do so. Because it is likely that these activities will appear to the user as irrelevant rather than necessary, it is unlikely that users will engage in explicit feedback activities. If users do not engage in relevance feedback activities, then there is no evidence upon which to provide personalization. Implicit relevance feedback techniques offer some solution to this problem.

Implicit relevance feedback techniques unobtrusively obtain information about users, their needs and document preferences, by watching their natural interactions with systems. Implicit feedback techniques are attractive methods for tailoring system responses to individual user preferences since they do not require additional effort on the part of the user. Some behaviors that have been most extensively investigated as sources of implicit feedback include reading time, saving, printing, selecting and referencing. The primary advantage to using implicit feedback techniques is that such techniques remove the cost to the user of providing explicit feedback. Implicit measures are generally thought to be less accurate than explicit measures (Nichols, 1997), but since large quantities of implicit data can be gathered at no extra cost to the user, they are thought to be attractive alternatives to explicit techniques.

The purpose of this chapter is to present a framework for understanding and studying how behavior can be used as implicit relevance feedback. This chapter begins with a classification and discussion of behaviors that have been used as implicit relevance feedback. This is followed by a general discussion and characterization of implicit feedback research. A selection of this research is then presented to illustrate how such studies have been conducted and how feedback has typically been measured and used. This chapter ends with a discussion of key issues and problems associated with implicit feedback and challenges for future research.

2. CLASSIFICATION OF BEHAVIORS USED FOR IMPLICIT FEEDBACK

Nichols (1997) provided the first classification of potential behavioral sources of implicit feedback by categorizing the actions that a user might be observed performing during information-seeking. For instance, reading time is considered as an action that allows one to "Examine" an information object. Nichols (1997) goes on to discuss the potential of, and the problems with, various sources of implicit feedback, as well as the costs. Specifically, Nichols (1997) states, "each implicit rating will probably contain less value than an explicit rating but the appropriate cost-benefit trade-off for different types of implicit data will have to be determined empirically." Oard and Kim (2001), Claypool et al. (2001), and Kelly and Teevan (2003) followed-up this initial classification with conceptual classifications of their own. Building on Nichols's (1997) classification, Oard and Kim (2001) proposed a framework of observable behaviors, which could be used to model interests using two axes, *Behavior Category* and *Minimum Scope*.