Modeling Information Content Using Observable Behavior

Douglas W. Oard and Jinmook Kim

Human Computer Interaction Laboratory
College of Information Studies and
Institute for Advanced Computer Studies
University of Maryland, College Park, MD 20742-4345
(oard,jinmook)@glue.umd.edu
phone: (301) 405 - 7590
fax: (301) 314 – 9145

Abstract

This paper presents a framework for modeling the content of information objects such as documents and video programs based on observation of how users interact with those objects in the course of information seeking and use. Four categories of potentially observable user behaviors are identified: examination, retention, reference, and annotation. The framework draws together techniques from the information filtering, Web searching and citation indexing, and identifies the natural scope (portion of an object, complete object, or collection of objects) at which each behavior can be observed. The process of using observations as a basis for identifying information that may be of interest to specific users is addressed briefly, and alternative system architectures are proposed. The paper concludes by identifying some open issues that could have significant implications for the utility of information content models that are based on observable behavior.

1. Introduction

Searchers are faced with the classic problem of finding a needle in a haystack, but the haystack is growing so rapidly that there is continual demand for improved search technology. Our fundamental approach to search has not changed, however. We still model the information needs of the user somehow (typically, using a query or a profile). We model the content of the information objects somehow (often by counting terms). We then use some technique (Boolean matching, vector space similarity, ...) to match the model of the need with the models of the available content. In some cases (e.g., relevance feedback or automatic profile learning), we base the model of the need in part on the model of the content. Improved models of information content could thus be used directly to improve any type of search, and in some cases could be leveraged to improve the model of the user's information needs. In this paper we examine one broad class of information content models, those based on observation of user behavior. Our goal is to provide a framework that integrates what has previously been seen as a diverse set of techniques into a coherent whole that can be used as a basis for design.

Information systems that represent the information content of an object on the basis of feedback provided by users of that object are now often referred to as *recommender systems* (CACM, 1997). We can consider observations of user behavior to be a form of *implicit feedback*, so the framework that we propose fits well with that perspective. Recommender systems are often associated with information filtering, a term used to describe so-called “push” services designed to find new information on pre-specified topics. The same representations can, of course, also be used by so-called “pull” services in which users seek to retrieve information from existing collections. Information filtering systems can be classified into two forms: content-based and collaborative (Oard, 1997). Recommender systems are typically contrasted with so-called “content-based” systems, in which information content is represented on the basis of the terms that they contain. These categories are not mutually exclusive, however. In this paper we focus exclusively on representing document content based on observable behavior. But, as Table 1 illustrates, the design of data structures that integrate both types of information would be quite straightforward. If users with similar interests are known (or can be discovered), the right side of the table might be useful. If terms related to the topic of interest are known, the left side might be useful. When both types of evidence are available, a combination-of-evidence strategy would be appropriate.

	Term Frequency					Reading Time			
	snails	**lizards**	**turtles**	**Birds**	**fish**	**User 1**	**User 2**	**User 3**	**User 4**
Doc 1	2				3	20		17	
Doc 2		5				3	22		
Doc 3			2		1			45	21
Doc 4		1		4	2		14	25	
Doc 5	1			1		37			

Table 1. Term frequency and reading time for five documents.

There is already good reason to believe that evidence from observable behavior can be useful. In information filtering research, for example, it is common to assemble a set of training instances that have been hand-labeled as relevant or as not relevant. Numerous studies have shown that such explicit feedback is clearly useful (e.g., Yan & Garcia-Molina, 1995; Goldberg et al., 1992). Similar results are routinely obtained for relevance feedback techniques (which also depend on explicit feedback) in information retrieval experiments. In practice, however, system developers are often frustrated to find that users of operational systems make little use of relevance feedback mechanisms. This probably results at least in part from the time and effort required to provide explicit feedback and from the additional cognitive load that may result from the need to decide on an explicit rating for each document. This experience suggests that over-reliance on explicit feedback would be unwise. In the next section, we review the emerging literature on implicit feedback in the context of information filtering. In Section 3 we introduce our framework for representing information content based on observable behavior, situating the research on implicit feedback in a broader context. Finally, in Section 4 we present some preliminary thoughts on how such information content representations might be effectively used in large-scale systems. Section 5 concludes the paper with some comments on future research directions suggested by the framework we have introduced.

2. Implicit Feedback: Lessons from Information Filtering

Implicit feedback may bear only an indirect relationship to the user's assessment of the usefulness of any individual document. But because it can be collected ubiquitously (and thus potentially in great quantities), the potential impact of implicit feedback might ultimately be even greater than that of explicit feedback. InfoScope, a system for filtering Internet discussion groups (USENET), utilized both implicit and explicit feedback for modeling users (Stevens, 1993). Three sources of implicit evidence were used: whether a message was read or ignored, whether it was saved or deleted, and whether or not a follow up message was posted. In summarizing this groundbreaking study, Stevens observed that implicit feedback was effective for tracking long-term interests because it operates constantly without being intrusive.

Morita and Shinoda (1994) introduced another source, proposing an information filtering technique based on observations of reading time. They conducted a user study over a six-week period with eight users to determine whether preference for USENET messages was reflected in the time spent reading those messages. The results showed a strong positive correlation between reading time and explicit feedback provided by those users. They also discovered that treating messages that the user read for more than 20 seconds as relevant actually produced better recall and precision in an information filtering simulation than would have been achieved using the messages explicitly rated by the user as relevant. Konstan et al. (1997) repeated this study in a more natural setting, distributing modified software that allowed volunteers to participate in a recommender system trial in which both explicit feedback and reading time were recorded for a small set of USENET discussion groups. Their results indicated that recommendations based on reading time could be nearly as accurate as recommendations based on explicit feedback. They also suggested some additional observable behaviors, including printing, forwarding, and replying privately to a message, as sources for implicit ratings.

Kim et al. (2000) conducted experiments with two user groups, examining how observations of reading time and printing behavior for journal articles might be used jointly to build a better user model than could be built using either source alone. They found that it was useful to condition the interpretation of reading time on the type of document (reading times for abstracts of relevant journal articles were typically longer than previous researchers had reported for relevant USENET news articles) and that printing behavior adds additional evidence beyond that which could be inferred from reading time alone. Not surprisingly, all of the documents that were printed were relevant, but the reading time for many of these documents was quite short.

Nichols (1997) sought to construct a comprehensive view of implicit feedback, with a focus on its use in information filtering systems. He presented a list of potentially observable behaviors; adding purchase, assess, repeated use, refer, mark, glimpse, associate, and query to those mentioned above. Oard and Kim (1998) extended that work, organizing observable behaviors into three broad categories: examination, retention, and reference. In the next section, we present a further refinement of that framework, adding an additional category (annotation) that results in unification of implicit and explicit feedback in a single framework.

3. A Framework for Observable Behavior

The framework summarized in Table 2 categorizes behaviors that may occur in the course of information seeking and use. We refer to these as "potentially" observable behaviors because the actual set of behaviors will naturally depend on the capabilities provided by the system and on the goals of the user. Furthermore, the observability of any specific behavior will depend upon the way in which the interface can be instrumented, considering both technical issues and social concerns such as privacy. The central (vertical) organizing principle of the framework is the purpose of the behavior, with a secondary (horizontal) distinction based on the minimum scope of the object being manipulated. Along the vertical axis, a user may *examine* a document, *retain* it, *refer* it to other people, and/or *annotate* it. Along the horizontal axis, a user may *view* a portion of a document, *select* an entire document, or *subscribe* to a collection of documents. The segment level includes operations whose minimum scope is a portion of an object (e.g., viewing a screen), the object level includes behaviors whose minimum scope is an entire object (e.g., purchase), and the collection level includes behaviors whose minimum scope includes more than one object (e.g., organization). By "minimum scope" we mean the smallest unit normally associated with the behavior – behaviors thus have analogues at larger scopes (e.g., viewing an entire document), but not normally at smaller scopes (e.g., purchasing a paragraph). The choice of "segment," "object" and "class" as labels is intended to be inclusive, since the ideas captured in the table generally apply equally well to non-text modalities such as video or music with only minor variations (e.g., listen rather than view).

Minimum Scope

Behavior Category	Segment	Object	Class
Examine	View Listen	Select	
Retain	Print	Bookmark Save Delete Purchase	Subscribe
Reference	Copy-and-paste Quote	Forward Reply Link Cite	
Annotate	Mark up	Rate Publish	Organize

Table 2. Potentially observable behaviors.

The *examine* category consists of *view, listen* and *select*. Information systems often provide brief summaries of several promising documents, so selection of individual objects for further examination can provide evidence regarding the user's assessment of those summaries. When applied to the Web, this type of evidence is often referred to as a "click stream." The minimum scope for viewing is a portion of a document, since a suitably instrumented interface could determine which lines are displayed at any time. Evidence at that scale could be used to construct rich usage histories that capture the same sort of information that well used books provide when they naturally fall open to a page that others have found to be useful (Hill, et al., 1992). Similarly, users of audio or video might choose to listen to only a portion of a program. Observing behavior at a scale below that of a complete object might provide more precise evidence of the user's intentions than object-scale observations alone, but at the cost of a somewhat more complex data collection effort.

The *retain* category groups behaviors that suggest some degree of intention to make future use of an object. *Bookmarking* a web page is a simple example of such a behavior. Rucker & Polanco (1997), for example, constructed a recommender system using bookmark lists. *Saving* the object itself is the obvious alternative, something Stevens (1993) used as a source of implicit feedback. The retention category is distinguished by the possibility of directly observing evidence of negative evaluations as well. When retention is a default condition, as in some electronic mail systems, a decision by the user to *delete* an object might support an inference that the deleted object is less valued than other objects that are retained. When information access is priced on a per-item basis, *purchase* decisions can offer extremely strong evidence of the value ascribed to an object. Similar information is available at a somewhat coarser scope when users purchase *subscription* access to certain types of content (e.g., subscription to a separately priced cable television channel). Retention behavior can also occur with a narrower scope. Although users often *print* entire documents, the cost of printing can provide an incentive to only print important parts of longer documents. It is actually somewhat artificial to place printing in a single category in our framework. We have grouped printing with retention because of the permanence of the printed page. Users sometimes print documents merely to facilitate examination, however, because paper still has many advantages over electronic displays. Printing also overlaps into the next category (reference), since users might sometimes print a document with the intention of forwarding it to another individual or using the printed document within another object (such as a poster). Nevertheless, printing is often associated with a desire for retention, so we find our grouping useful.

Each activity in the *reference* category has the effect of establishing some form of link between two objects. Refer behaviors can be represented as a graph with directed edges. A segment is the minimum scope for the origin of any edge, but the destination of an edge is a complete object for some behaviors and a segment for others. *Forwarding* a message, for example, is an object-scale action that establishes a link from the original message to the new one. Similarly, *replying* individually or posting a follow up message to some form of group venue such as a mailing list establishes the same sort of link. Goldberg et al. (1992) described a simple example of this in which users could construct an electronic mail filter to display messages that their colleagues had taken the time to reply to. Hypertext links from one Web page to another and bibliographic citations in academic papers also create links from a portion of an object (characterized, perhaps, by some neighborhood around the link itself) to an entire object.[1] The PageRank algorithm used in Google provides an obvious example of how useful hypertext links can be (Brin & Page, 1998). Garfield (1979) used bibliographic *citations* as a basis for retrieval, and that technique has recently been applied to the Web as well (Lawrence, et al., 1999). At the scope of a segment, selective inclusion of a portion of one document in another, using either *copy-and-paste* or a manually retyped *quotation* (that might be automatically detected and associated with the original passage), creates a link from a portion of one information object to a portion of another.

Observable behaviors in our fourth category, *annotation*, are actions that intentionally add to the value of an information object. At the segment scale, a user may *mark-up* a portion of a document by highlighting a passage or attaching the electronic equivalent of a sticky-note if the system provides those capabilities. At the object scale, some entity might assign an explicit *rating* to the object for the benefit of other users. This is commonly done for journal articles during the peer review process, for example. Eventually, some entity might act to increase the accessibility of the object by *publishing* it, thereby creating an implicit endorsement of the value of the published information (Wang & Soergel, 1998). At the collection scale, a user might *organize* a set of objects to facilitate later access, thereby providing additional evidence of their value.

Our goal in presenting this framework is to provide developers with a useful perspective on the broad range of observable behaviors that might be used as a basis for representing information content. Some caveats are clearly needed, however. Our goal was to be descriptive rather than exhaustive, so we have based the assignment of behaviors to categories on our intuition about typical conditions encountered during information access and use. Some adjustments might therefore be needed to apply these categories to specific applications (e.g., if users were able to bookmark segments of documents). Furthermore, although we have tried to choose terms that apply equally well to modalities other than text, the vast majority of the work we have drawn on has been focused on text. As we gain more experience with other modalities, it may become clear that we have overlooked some important categories of behaviors (perhaps humming a tune could be taken as an implicit endorsement of a song?). Nevertheless, we believe that the framework shown in Table 2 accurately reflects what we know today. In the next section, we outline some ways in which information content representations might be constructed by observing these behaviors.

[1] In both cases, segment-scale variants of the behavior exist, but are less commonly used.

4. Modeling Information Content

The nature of an appropriate representation of information content is intimately related to the way in which the user's information need is represented and the available techniques for matching. In this section, we use information filtering as an example, and restrict our attention to recommendation-based approaches. Such systems are often referred to as “collaborative filtering” systems. In collaborative filtering systems, an initial user needs profile is typically constructed manually, often by selecting from a set of predefined stereotypes. The behavior of some users with respect to some information objects is then observed, and those observations are used to predict the value of those same information objects to other users. Figure 1 shows two simple strategies for accomplishing this: inference followed by prediction, or prediction followed by inference. In the “Inference-Prediction” strategy, the inference stage seeks to produce ratings similar to those that a user would have explicitly assigned, and then the prediction stage then uses those estimated ratings to predict the desirability of the object to another user. Konstan et al. (1997) adopted this perspective when evaluating how well observed reading time predicted explicit ratings for individual articles. The “Prediction-Inference” strategy, by contrast, uses observations of the behavior of some users to predict the behavior of some other user, and then the inference stage seeks to estimate the value of the information based on the predicted behavior. Stevens (1993) implemented a variant of this strategy, predicting examination duration for a new USENET news article based on the examination durations for similar articles in the past and then constructed content-based queries that would select articles with long predicted examination durations. This equates to a degenerate inference stage in which desirability is assumed to increase monotonically with examination duration.

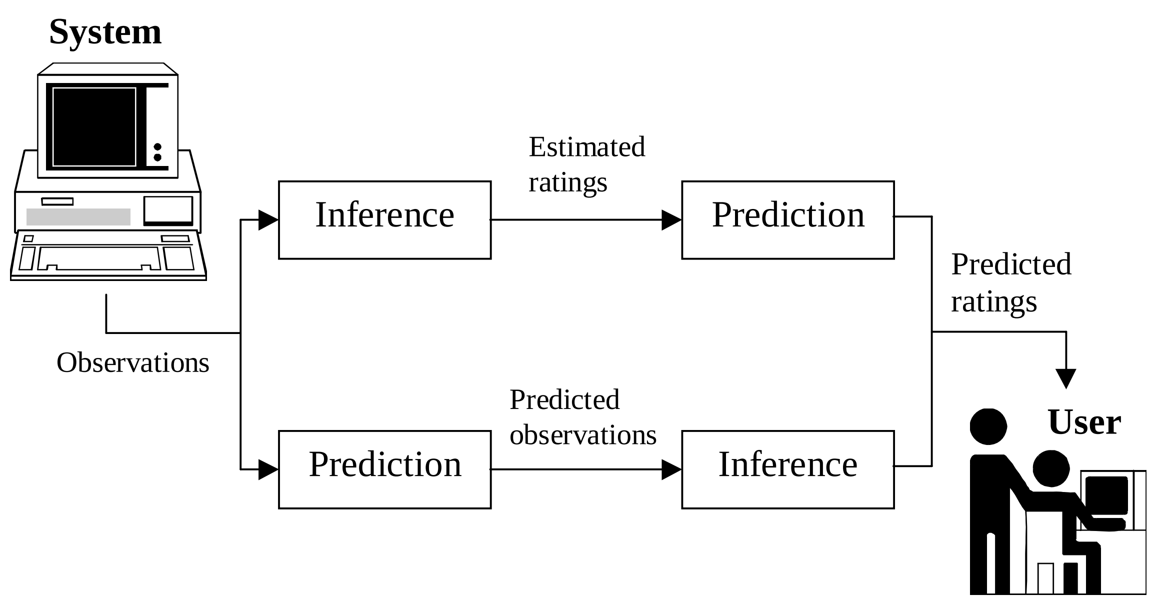

Figure 1. Alternative strategies for making recommendations

Collaborative filtering systems based on the second strategy might be more flexible, since participating users might draw different inferences from the same observations if they did not share a common set of objectives. On the other hand, collaborative filtering systems using the first strategy would likely have more context available locally for interpreting observations than would be available at other points in the network. In small-scale applications, it may be possible to perform all computations on a single server. But scaling these techniques to a global network will clearly call for a distributed implementation in which the context needed to support inference exists partially at one server and partially at another. In such cases, it might be worth considering a hybrid approach in which some preliminary interpretation is performed locally when the observation is made, additional changes might be made at other points in the network (perhaps to conserve bandwidth through aggregation), and then the ultimate inference is performed in the user's server.

5. Conclusion and Future Work

Users of information systems exhibit a range of behaviors that might be observable in one application or another. Our goal in this paper has been to draw together research on this topic from diverse traditions into a coherent framework that can guide practice. We have illustrated the potential of these techniques with citations to the literature on collaborative filtering, citation indexing, and Web search. Several potentially observable behaviors have been shown to be useful in one context or another, and some of those behaviors (e.g., total reading time) are easily observed using widely deployed Web browsers and servers. Observing other types of behavior (e.g., the distribution of reading time within a document) would require new capabilities that the standard Web infrastructure does not yet provide. Another important limitation of our present understanding is that relatively few studies have examined combination of evidence from more than one type of observable behavior. The combination of behavior-based and content-based evidence is also an important issue, and one that has only begun to be explored. In the last section we offered some ideas on the use of observable behavior for content modeling in large-scale systems. If information content models based on observable behavior are to reach their full potential, such system architecture issues will also need attention. Finally, a number of social issues must be addressed before the full potential of this technology can be realized. Foremost among these is privacy protection for the users whose behavior we observe, and for the users that make use of the information content models. Another important issue is informed consent; how can we explain to a user the use that we will make of the information we obtain by observing them? Although a number of challenges remain, we are confident that information content models based on observable behavior offer substantial potential to improve the search effectiveness in a globally networked environment.

References

Brin, S. and Page, L. (1998) The anatomy of a large-scale hypertextual Web search engine. *Coputer Networks* 30(1-2) 107-117.

CACM (1997) Special Issue on Recommender Systems, *Communications of the ACM*, 40(3), March.

Garfield, E. (1979*) Citation indexing: Its theory and application in science, technology, and humanities*. New York: Wiley-Interscience.

Goldberg, D., Nichols, D., Oki, B. M, and Terry, D. (1992) Using collaborative filtering to weave an information Tapestry. *Communication of the ACM*, December, 35(12), 61-70.

Hill, W. C., Hollan, J. D., Wroblewski, D., and McCandless, T. (1992) Edit wear and read wear. In *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI'92)*, 3-9.

Kim, J., Oard, D. W., and Romanik, K. (2000) User modeling for information access based on implicit feedback. Technical Report: HCIL-TR-2000-11/UMIACS-TR-2000-29/CS-TR-4136, University of Maryland, College Park. ftp://ftp.cs.umd.edu/pub/hcil/Reports-Abstracts-Bibliography/2000-11html/2000-11.html

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., and Riedl, J. (1997) GroupLens: Applying collaborative filtering to Usenet News. *Communication of the ACM*, March, 40(3), 77-87.

Lawrence, S., Bolacker, K. and Giles, C.L. (1999) Indexing and Retrieval of Scientific Literature. In *Eighth International Conference on Information and Knowledge Management (CIKM'99)*, 139-146. http://www.neci.nec.com/~lawrence/papers.html

Morita, M and Shinoda, Y. (1994) Information filtering based on user behavior analysis and best match text retrieval. In *Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval*, pp. 272-281.

Nichols, D. M. (1997) Implicit ratings and filtering. In *Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering*, Budapaest, Hungary 10-12, ERCIM. http://www.ercim.org/publication/ws-proceedings/DELOS5/index.html

Oard, D. W. (1997) The state of the art in text filtering. *User Modeling and User-Adapted Interaction*, 7(3), 141-178.

Oard, D.W., and Kim, J. (1998) Implicit Feedback for Recommender Systems. In *AAAI Workshop on Recommender Systems*, Madison, WI: 81-83. http://www.glue.umd.edu/~oard/research.html

Rucker, J. and Polanco, M. J. (1997) Personalized Navigation for the Web. *Communications of the ACM*, March, 40(3): 73-89.

Stevens, C. (1993) *Knowledge-based assistance for accessing large, poorly structured information spaces*. Ph.D. thesis, University of Colorado, Department of Computer Science, Boulder. http://www.holodeck.com/curt/mypapers.html

Wang, P. and Soergel, D. (1998), A cognitive model of document use during a research project. *Journal of the American Society for Information Science*, 49(2): 115-133.

Yan, T.W. and Garcia-Molina, H. (1995) SIFT – A tool for wide-area information dissemination. In *Proceedings of the 1995 USENIX Technical Conference*: 177-186. ftp://db.stanford.edu/pub/yan/1994/sift.ps.