A User Modeling System for Personalized Interaction and Tailored Retrieval in Interactive IR

Diane Kelly and Nicholas J. Belkin

School of Communication, Information and Library Studies, Rutgers University, 4 Huntington St., New Brunswick, NJ 08901. E-mail: diane@scils.rutgers.edu

We present a user modeling system for personalized interaction and tailored retrieval that (1) tracks interactions over time, (2) represents multiple information needs, both short and long term, (3) allows for changes in information needs over time, (4) acquires and updates the user model automatically, without explicit assistance from the user, and (5) accounts for contextual factors such as topic familiarity and endurance of need. The proposed system contains three major classes of models: general behavioral, personal behavioral and topical. The general behavioral model describes how information search and use behavior can be used to identify and track information needs. The personal behavioral model characterizes an individual user's information search and use behavior with regard to document preference and states of knowledge. Finally, the topical model characterizes the user's information seeking needs. We describe how such a model can be used to personalize search interactions and tailor system responses to individuals across multiple information seeking sessions.

Introduction

As information becomes increasingly available, users are faced with an overabundance of sources in which they must choose. Collections are no longer homogeneous sets of documents with conventional structure and standard vocabulary, but are instead, heterogeneous sets of documents with varying structure and undifferentiated vocabularies. Not only is information increasingly available, but it is also increasingly accessible. As more information is distributed electronically, a user's information seeking activities are no longer bound geographically or temporally. Now, more than ever, it is critical for systems to obtain a more accurate representation of the user's information needs, document preferences and states of knowledge and to maintain these representations over time.

Most IR systems assume that information seeking episodes are discrete and unrelated. Recent work (Lin & Belkin, 2000; Spink, 1996; Spink, Griesdorf & Bateman, 1999) has challenged this assumption and provided evidence that information seeking typically occurs across multiple search sessions and that information needs often evolve throughout the process of successive searching. Traditional IR systems have construed searching to occur within a single search session. Search sessions and users are synonymous terms; for each search session, the system assumes a new user. Because of this, much information about the user and the user's preferences and states of knowledge is lost. Some systems may provide the illusion of continuity, such as a query history or saved documents lists, but this information only acts as a memory aid or navigation feature. The system does not use this information to characterize the user and the user's preferences or to aid in the retrieval of documents. Each time the user initiates searching, he/she must begin anew as far as the IR system is concerned. Even though the user's information seeking activities can occur during multiple search sessions across multiple time periods, the system views these activities as occurring within specific instances of time. Essentially, the system knows nothing about the user and is unable to determine that the user searching currently for information about "The Flea" is the same user that searched for information about the poetry of John Donne just four days ago. To know even this small bit of information would make the difference between returning documents about flea control and flea markets and returning documents about the poem.

User Modeling

User modeling (UM) offers the potential of individuating users and tracking the information seeking behavior and information needs of the user over time. Generally speaking, a user model is a description of a user, created or selected by the system, that facilitates interaction between the two (Allen, 1997). One of the underlying goals of UM is to predict user responses, thereby creating more effective, efficient and personalized interactions by tailoring system responses to individual user preferences. Systems that incorporate user models have been traditionally termed "intelligent systems" because a representation of the user is created by the system and used to reason about the user and

Kelly, D. & Belkin, N. J. (2002). A user modeling system for personalized interaction and tailored retrieval in interactive IR. In *Proceedings of Annual Conference of the American Society for Information Science and Technology (ASIST '02)*, Philadelphia, PA, 316-325.

the user's document preferences. The system then uses the information to personalize interactions.

Knowledge Elicitation

In order to build a model of the user, the modeling function must somehow obtain information from the user about his/her interests and preferences. Human intermediaries are quite successful at eliciting knowledge from users about information needs and using this information to personalize interactions (Belkin, 1984; Belkin, Brooks, & Daniels, 1987; Daniels, 1987; Ingwersen, 1982; Saracevic, Spink, & Wu, 1997; Spink, Goodrum, Robins, & Wu, 1996). Belkin (1984) finds that intermediaries elicit information from the user in order to characterize the user's information problem and to determine appropriate retrieval strategies. Daniels (1987) finds that intermediaries elicit knowledge from users with respect to goals, knowledge of the topic and task and familiarity with IR systems. This body of work also suggests that the intermediary dynamically updates much of the information that is elicited from users throughout the interaction. Saracevic, Spink & Wu (1997) identify shifts that occur during the intermediary-user interaction. These shifts reflect things such as changes in the user's states of knowledge, specifically with regard to understanding and learning, demonstrating that UM techniques for interactive IR should be iterative and long term and should model things such as the user's evolving states of knowledge.

1.1.1 Explicit Techniques

Knowledge elicitation in interactive IR and information filtering (IF) has proven to be a significant challenge. Attempts have been made in IR to develop systems that simulate the user modeling functions performed by typical human search intermediaries (Brajnik, Guida & Tasso, 1987; Croft & Thompson, 1987; Oddy, 1977; Rich, 1983; Vickery & Brooks, 1987). One of the underlying goals of UM in IR is to create a representation of the user and use this representation to select appropriate retrieval strategies and information objects. For instance, Grundy (Rich, 1983) employs UM to recommend works of fiction to users; I^3R (Croft & Thompson, 1987) employs UM to retrieve documents for users.

Knowledge elicitation techniques that have been employed for user modeling in IR and IF have typically required the user to explicitly participate in the process of model construction and maintenance. It is often the case that the user is asked to specify keywords or to select and mark the relevancy of documents. In other cases, users are required to answer a series of questions about their interests. Explicit techniques for model acquisition are problematic because they are cumbersome and labor-intensive and they assume that users can articulate or identify adequately their information needs. Requiring the user to identify keywords that characterize the information need constrains the user and the user's need in a variety of ways. When a user approaches an IR or IF system, it may be the case that the user is unsure of his/her interests (Belkin, 2000). In IR systems, users are often unable to clearly articulate and describe their information needs (Belkin, Oddy & Brooks, 1982). If the user is just beginning to learn about a specific topic, the user may lack the vocabulary to describe the topic. Indeed, a user's ability to expand his/her vocabulary and knowledge of the topic may be impeded if the only documents returned are those that contain words, of which the user is already aware. Additionally, the user may not be able to provide enough terms to accurately or exhaustively cover the scope of his/her information need.

Most IR and IF systems only accommodate users who have a single information need. These systems do not accommodate a single user who has multiple, unrelated needs. There is no consideration for how these multiple needs might be integrated into a single user model or of how the representations of these various different needs might affect retrieval. Furthermore, information needs often change over time. Recent work by Lam & Mostafa (2001) and Widyantoro, Ioerger & Yen (2001) has attempted to detect and track changes in the user's interests over time. However, accounting for changes in the user's familiarity and/or learning has remained relatively unexplored. The user is assumed to have the same understanding of an information need if he/she searches one time for the need or twenty times. While in some cases users can manually edit the contents of their user model, it does not seem realistic to assume that users will make an effort, or even remember to update and maintain their user models over time.

Another problem with IR and IF modeling systems is that users are required to read and rate a large number of documents in order for the user model to be constructed. Clearly, it is unrealistic to believe that users have the extra time required to read and rate a large number of articles everyday. One can imagine that the extra time required to rate each article might prevent users from viewing additional articles of interest. Ultimately, the cognitive burden placed on the user by requiring him or her to rate such a large number of articles is problematic. The cost of having to read and rate a large number of documents is clear to the user while the benefit is not. It is clear to the user that he/she will have to spend additional time engaged in an activity which does not appear to be immediately addressing his/her primary needs; it is not clear how this will help. Unless the benefit is obvious to users, documents will remain unread and un-rated. Without information about user preferences, it is difficult for systems to create personalized user models.

1.1.2 Implicit Techniques

There have been a number of proposals of methods for obtaining information about user preference implicitly. Implicit techniques unobtrusively obtain data about user preference through the observation of "normal" human behavior during information search and use activities. Implicit techniques appear to be attractive candidates for tailoring system responses to individual user preference without requiring the user's explicit participation. Advantages to using implicit methods are that they (1) remove the costs of having a user read and rate a large number of documents and (2) are easy to obtain. Implicit measures are generally thought to be less accurate than explicit measure (Nichols, 1997), but they can be gathered at no cost to the user and in large quantities, and can be combined to obtain a more accurate representation of user interests. Implicit measures can also be used to supplement a smaller number of explicit ratings.

Implicit feedback techniques have been primarily investigated in IF and recommendation systems. Behaviors most extensively investigated as sources for implicit feedback have been selection, reading, saving and printing. For instance, Morita and Shinoda (1994) found a strong tendency for users to spend a greater length of time reading those articles rated as interesting, as opposed to those rated as not interesting. Others have replicated this finding in similar environments (Billsus & Pazzani, 1999; Konstan, Miller, Maltz & Herlocker, 1997; Seo & Zhang, 2000). Other behaviors that have been explored as sources of implicit feedback include scrolling and bookmarking (Billsus, Pazzani & Chen, 2000; Kamba, Sakagami, & Koseki, 1997; Lieberman, 1995; Mladenic, 1999; Oard & Kim, 1998; Rucker & Polanco, 1997). Oard & Kim (2001) & Nichols (1997) provide conceptual classifications of potential behavioral sources of implicit feedback.

While implicit techniques offer substantial promise with regard to capturing and modeling user preference automatically, they are not without problems. Most of the techniques explored have been a secondary focus of the research in which they are reported. Once the general relationship is discovered, it is applied to all users without regard to individual behaviors. If the goal of user modeling is to tailor system responses to individual user preference, then research clearly needs to focus on not only identifying valid implicit measures of user preference at the general level, but also on determining if and how these general relationships hold for individual users. In the interactive IR community, it is no secret that different users exhibit different searching behaviors. To discover a valid measure for document preference based on groups of users, and then apply it to all users, is to do little more than create user models based on canonical or stereotypical users.

One of the most extensively used behaviors for implicit modeling, reading time, has been discovered recently to be not as valid and reliable a measure as once thought. While this measure has been applied in numerous systems (Morita & Shinoda, 1994; Konstan, et. al, 1997; Billsus & Pazzani, 1999, Seo & Zhang, 2000), Kelly & Belkin (2001, 2002) find that several factors, such as specific task, topic and user characteristics, confound the general finding in complex ways. Even though Kelly & Belkin (2002) were able to replicate the findings of Morita & Shinoda (1994) and others (Konstan, et. al, 1997) with regard to reading time, a closer inspection of the data demonstrated that the significant relationship held for less than one-third of the study participants and that for three participants, the relationship was in the opposite direction from that which was expected. Kelly & Belkin (2002) were unable to construct a time-based rule based on the significant results which could predict the relevance of documents within the same data set sufficiently and accurately. These results demonstrate the danger of using a single measure based on the behavior of multiple users as evidence for the creation of individual user models. The authors suggest that combining evidence from multiple measures may be a more valid and reliable approach.

The rate of occurrence of the behaviors may be affected by specific contextual factors, such as topic familiarity, type of information need, type of task and searching environment. As a user learns more about a particular topic, his/her information search and use behavior with respect to that topic may change. For instance, as a user becomes more familiar with a topic, the user may be able to more quickly judge the relevance of documents or the user may save documents less frequently. Kelly & Cool (2002) have found evidence that indicates that as one's topic familiarity increases, one's reading time for retrieved documents decreases. It may be the case that changes in a user's information search and use behaviors reflect changes in the user's state of knowledge about a particular topic.

The application of implicit techniques to user modeling is problematic because it occurs without regard to individual differences. It may be the case that certain users exhibit only certain behaviors. It may be the case that a combination of evidence from multiple sources provides one with a more accurate model of individual preference. Finally, it may be the case that behaviors are only valid for certain users if they co-occur with other behaviors. Thus, a significant challenge is the development of methods for the application and adaptation of general relationships between various sources of implicit feedback to individual user model construction. If the behavioral model is invalid and unreliable at the individual level, then so too will be the resulting model of the user's information needs.

A User Modeling System for Personalized Interaction and Tailored Retrieval

Given the evidence cited above, we suggest the following desiderata for an ideal UM system for interactive IR. A UM system for personalized interaction in IR should:

- Track information needs and interactions over time.

- Represent multiple information needs, both short and long term.
- Allow for changes in information needs over time.
- Acquire and update the user model automatically, without explicit assistance from the user.
- Account for differences such as topic familiarity and endurance of need.

Finally, a UM system for personalized interaction in IR should do just that, *personalize* interaction.

We propose that such a system should contain three major classes of models: general behavioral, personal behavioral and topical. The general behavioral model functions to describe how information search and use behavior can be used to identify and track information needs. The characterizations found in this model are general to all users. The personal behavioral model functions to characterize an individual user's information search and use behavior with regard to document preference and states of knowledge. The characterizations found in this model are specific to the individual user. Finally, the topical models function to characterize the user's information seeking needs. Like the personal behavior model, these characterizations are specific to the individual user. Both types of behavioral model include some aspects of context, such as topic familiarity and endurance of need. The topical model includes a representation of the terms and concepts that are associated with a particular information need. It is proposed that topical models be inferred through characterizations defined by the behavioral models that describe the relationship between a user's information seeking behavior and information needs. For the purposes of this paper, states of knowledge can be thought of as the concepts that a user associates with a particular topic, how these concepts are related to one another and how groups of concepts relate to, or cluster with, other groups of concepts. Document preference simply refers to a user's choice for one document over another. It is surmised that this choice reflects states of knowledge, as well as the user's interests in a topic. Figure 1 provides a brief outline of how such a system might work.

The General Behavioral Model

The General Behavioral Model (GBM) functions to describe how implicit feedback can be used to identify and track a user's information needs. The GBM contains characterizations of information search and use behavior that can be used as starting points upon which to begin the construction of personalized user models. We suggest that the behavioral characterizations found in the GBM be acquired through a longitudinal study of the information search and use behaviors of a large group of users. The relationships described by the GBM can function as hypotheses that are used by the system to reason about an individual user's behavior and the relationship of this behavior to the user's document preference and states of knowledge during the construction and maintenance of personalized user models. The GBM can assist in the maintenance of personalized user models by offering possible explanations for previously unseen behavior exhibited by the user, changes in the user's behavior and/or irregular behavior exhibited by the user. In order to achieve these goals, the GBM might suggest a range of values for particular behaviors based upon the frequency of occurrence and/or co-occurrence of particular behaviors. The GBM might also achieve these goals by generating equations which describe the user's information search and use behaviors.

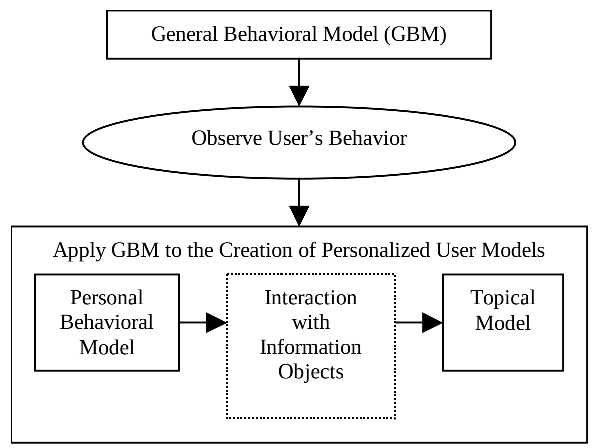

Figure 1. A UM system for personalized interaction

Personalized Behavioral Models

The personalized behavioral models are constructed iteratively through observation of the user's behavior and application and adaptation of the GBM to a specific user's behavior. When a new user begins searching, the GBM provides a mechanism in which to evaluate the user's behavior. The GBM identifies particular relationships and suggests possible values for observed behaviors. Again, these characterizations are proposed to assist with personalized user model construction by suggesting possible explanations for observable behaviors. It is expected that the relationships suggested by the GBM between information search and use behaviors, document preference and states of knowledge will be iteratively modified and adapted at the individual level as the system learns more about the user through observations that occur across time.

The personalized behavioral model created by the system may be similar to the GBM, or it may be quite different. It will most likely be the case that behaviors suggested by the GBM are exhibited by different users with different frequencies and regularities. One user may read faster than another user. For a certain user, a particular behavior may be strongly related to document preference, while for another user the behavior may be unrelated. It may be the

Kelly, D. & Belkin, N. J. (2002). A user modeling system for personalized interaction and tailored retrieval in interactive IR. In *Proceedings of Annual Conference of the American Society for Information Science and Technology (ASIST '02)*, Philadelphia, PA, 316-325.

case that behaviors are only useful in combination with one another or if they co-occur temporally with other behaviors.

One possible way that individual differences can be accounted for is through the association of weights with the occurrence of various behaviors. For instance, a set of behaviors may be initiated with a weight of 1. As the system learns more about the user, the weights associated with these individual behaviors can be modified accordingly. If the user rarely exhibits a certain behavior, then its weight can be progressively downgraded to 0. This approach can also be used for behavioral model maintenance. If the frequency of a particular behavior that was once identified as a useful source of a user's document preference decays over time, then the weight associated with this behavior can be altered. Alternatively, if a previously inhibited behavior begins to be exhibited by the user with some regularity, then the weight associated with this behavior can be altered as well.

Two hypothetical personalized behavioral models for User A and User B are displayed in Figure 2. For User A, query history, document viewing, bookmarking, bookmark structure and saving are all behaviors that have been identified as valid sources of implicit feedback for document preference and states of knowledge. These are all behaviors that User A exhibits with some regularity. For User B, the valid sources of implicit feedback for document preference and states of knowledge are a bit different. User B does not organize his/her bookmarks; thus, bookmark structure is not a valid source of evidence for document preference and states of knowledge. Note for User B, document viewing and bookmarking are only valid sources of evidence if they co-occur. If these behaviors do not co-occur, then they become only weak evidence from which to infer User B's document preference and states of knowledge. Figure 2 illustrates how personalized behavioral models might differ from user to user.

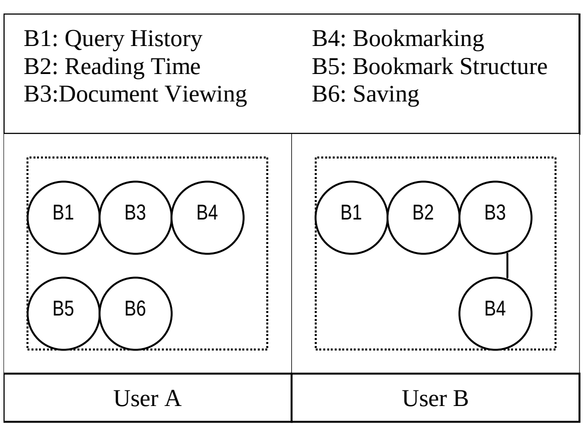

Figure 2. Hypothetical personalized behavioral models

Accounting for Contextual Factors

It is important for the GBM to identify and account for contextual factors, such as topic familiarity and endurance of need, and to describe how this affects the information search and use behaviors that the user exhibits. Some behaviors may only be exhibited by a user in certain searching environments or with certain types of tasks or topics. We are primarily concerned with two dimensions of information needs, level of familiarity and endurance of need. Level of familiarity reflects the user's knowledge of a topic. Level of familiarity might be measured according to amount of time spent searching for a particular topic. Endurance of need reflects that length of time that a user will be engaged with an information need. Most information needs have some temporal constraint associated with them. Some information needs may be confined to a particular time period and may have a specific, fixed date of termination. Other information needs may persist over time, with no clear boundary of use or date of termination. Endurance of need might be measured according to frequency of search for a particular topic. The relationship between behavioral models and these two dimensions of information needs is displayed in Figure 3.

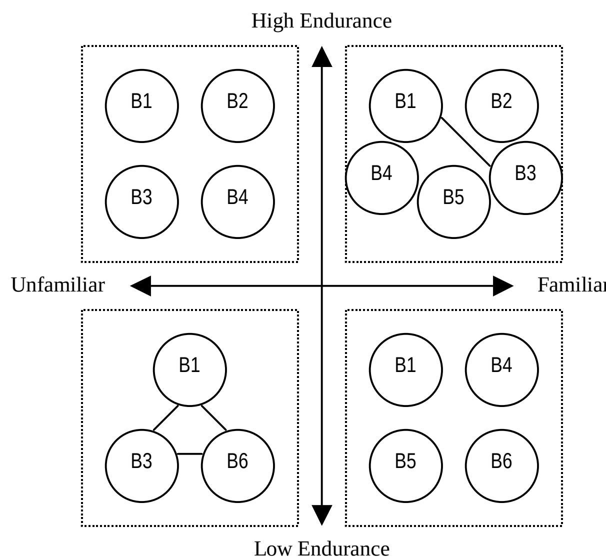

Figure 3. Relationship between behavioral models, level of familiarity and endurance of information need

The GBM functions to suggest possible methods for dealing with behavioral variations that individual users exhibit through the provision of alternative ranges of values for behaviors according to level of familiarity and endurance of information need. One can imagine different values for various behaviors based upon these two dimensions. For instance, it is likely that collections searched is a useful source of implicit feedback of user preference for an information need with high endurance; it is unlikely to be a useful source of user preference for an information need with low endurance. Of course, some low endurance needs

Kelly, D. & Belkin, N. J. (2002). A user modeling system for personalized interaction and tailored retrieval in interactive IR. In *Proceedings of Annual Conference of the American Society for Information Science and Technology (ASIST '02)*, Philadelphia, PA, 316-325.

may eventually turn into high endurance needs and the behavioral model may adapt accordingly. When a user initiates a search with a query that the system is unable to recognize and classify as high endurance need, the system can initialize the values that characterize low endurance needs that are given by the GBM. As the system learns more about the user and how the user behaves with low endurance needs, a personalized behavior model for low endurance needs for that particular user can evolve. If the user continues to search for this need over time, then it may be the case that this user's information search and use behaviors will begin to more closely resemble the type of behaviors that are exhibited by the user during information seeking of high endurance need. Consequently, the behavioral model that represents the information seeking behavior that is associated with high endurance needs for that particular user can be instantiated as the user's behavior changes over time.

Degree of familiarity with the topic can also affect the types of information search and use behaviors exhibited by the user. For instance, the relationships described by the GBM model might differ depending upon level of familiarity with a topic. Consider the relationship between reading time of relevant and non-relevant documents and topic familiarity. If a topic is new to the user, then it is likely that the user is unsure of the characteristics that distinguish relevant from non-relevant documents. When this occurs, reading time for relevant and non-relevant documents may be similar. Conversely, if the user is familiar with the topic, then it is likely that the user is better able to distinguish between relevant and non-relevant documents. When this occurs, reading time for relevant and non-relevant documents may be different. In this situation, reading time is useful, but in two very different ways. The GBM can account for variations in behaviors due to topic familiarity by suggesting a range of values based on degree of familiarity.

Table 1 displays an example of the range of values that the GBM might contain for reading time of relevant and non-relevant documents according to familiarity. Topic familiarity might be defined as a function of time; the longer amount of time that a user engages in information search and use behaviors about a given topic, the more familiar the topic becomes. The example in Table 1 identifies three degrees of familiarity based on observations made during the construction of the GBM, high familiarity (3); medium familiarity (2); and low familiarity (1). The identification of these levels is based upon the observation of three distinct relationships between reading time for relevant and non-relevant documents. It is important to note that Table 1 functions *only* as an example of how the GBM might derive a range of values for topic familiarity based on observed behavior; it is unclear if three levels of familiarity would emerge and/or if this range of values is even valid. Continuing with the example, Table 1 illustrates how reading time varies for relevant and non-relevant documents with respect to familiarity; as the user becomes more familiar with the topic, reading time for both types of documents decreases.

Table 1. Range of possible values for reading time and level of topic familiarity

Level of Familiarity	Relevant	Not Relevant
3 (high)	12 seconds	5 seconds
2 (medium)	17 seconds	14 seconds
1 (low)	20 seconds	25 seconds

The range of values suggested by the GBM for reading time can be used during the initial construction of the personalized behavioral models. Figure 4 demonstrates how this might occur using reading time as implicit feedback. In this example, User A initiates searching by entering a query. For our purposes, we will assume that User A is a recognized user of the system and the system contains a personalized user model for User A. The query that is issued by User A is unrecognized by the system. That is, the query cannot be matched to any of the pre-existing topic models that the system holds of the user. Thus, the system assumes that the information need that is represented by this query has low endurance and/or that the user's familiarity with the topic is low.

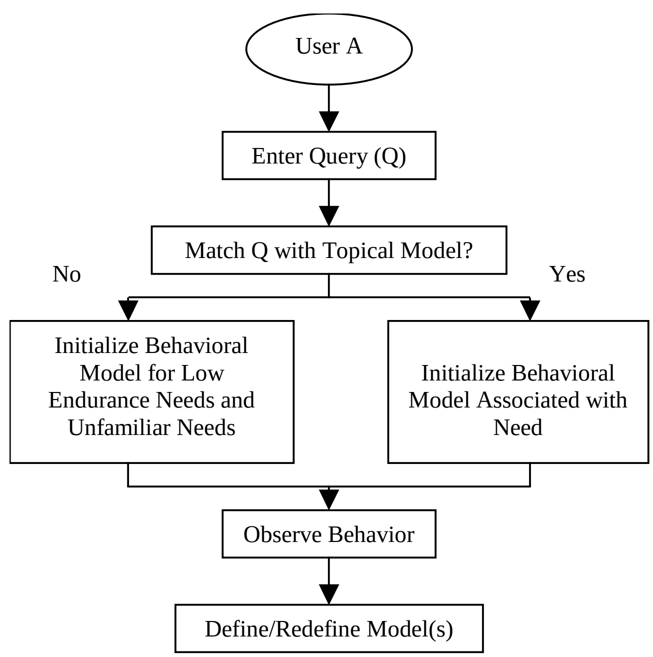

Figure 4. Personalization of interaction

Let us assume that a personalized behavioral model characterizing User A's behavior for unfamiliar topics has not been created during past interactions, but that a behavioral model characterizing User A's behavior for low endurance needs has been developed through past interactions. Table 2 displays User A's behavioral model for low endurance needs, as well as some other models of reading time that the system might have for User A. During the initial interactions, the system's observations and interpretations of the user's information search and use behavior is biased by User A's behavior model for low endurance needs and by the GBM's values for unfamiliar topics. At this point in the search session, it is unclear which model should be initialized. Now, suppose that the system discovers through observing the user's behavior that the user spends 11 seconds reading those documents that he/she identifies as relevant and 9 seconds reading those documents that he/she does not identify as relevant. In this example, the observed behavior most closely resembles that characterized by the User A's personalized behavioral model for low endurance needs.

Table 2. Models of reading time (RT) in seconds, of relevant (R) and non-relevant (NR) documents for User A.

Low Endurance Needs			High Endurance Needs			High Familiarity		
	R	NR		R	NR		R	NR
RT	11	9	RT	20	12	RT	13	5

Other examples of possible observed reading times for User A and the models that they activate are displayed in Table 3. This table is used in conjunction with Table 1 and Figure 4. Observing reading times of 11 seconds for relevant documents and 9 seconds for non-relevant documents would be matched to User A's preexisting behavioral model for low endurance needs, while observing reading times of 20 seconds for relevant documents and 25 seconds for non-relevant documents would not be matched to any pre-existing personalized behavioral model but would be matched to the low familiarity model found in the GBM. Since User A does not have a previously defined behavioral model for low familiarity, observing reading times of 20 and 25 seconds would initiate the construction of this model based on the values found in the GBM.

Table 3. Examples of reading time (RT) in seconds, observations and initiated action (BM: Behavioral Model).

RT Observations		Matches BM?	Matches GBM?	Action
R	NR			
11	9	**YES**	-	Initialize Low Endurance Model
		No		
13	5	**YES**	-	Initialize High Familiarity Model
		No		
20	22	Yes	**YES**	Define Low Familiarity Model
		NO	No	

Topical Models

The next section of this paper is concerned with the development of topical models. As mentioned earlier, personalized user models will consist of two classes of models: behavioral and topical. Topical models are the primary mechanism for tailoring retrieval. Topical models represent the user's information needs as collections of concepts. Topical models can be thought of as an associative network of concepts; relationships between concepts are represented through connections. Concepts are represented by clusters of terms that have been selected through the observation and monitoring of the user's information search and use behavior. For instance, terms might be extracted from the user's query. Terms might also be extracted from documents that the user has identified as useful in some way, such as when the user saves or prints a document. Behavioral models specify when and how topical information can be implicitly gathered from an individual's information search and use behavior. Thus, topical models are inferred based on the user's information search and use behavior. Personalized behavioral models provide methods for making these inferences.

Topical models can be initialized when a user initiates an information-seeking episode. When the user enters query terms that are associated with concepts found within a particular topical model, the topical model is initialized. Since the development of the topical models is based upon the user's past information search and use behaviors, topical models are associated with one or more behavioral models. Consequently, when a topical model is initialized, so too are one or more behavioral models. These behavioral models, in turn, will guide the system's observations and interpretations of the search behavior that the user exhibits during this particular information-seeking episode. For instance, consider a user who enters a query that is matched to Topical Model 5 (TM5). If TM5 exists, then it has some history of information seeking interactions associated with it. In other words, previous information search and use behavior has been associated with this model. Figure 6 is an extension of Figure 3. Figure 6 depicts the relationship between behavioral models, topical models, endurance of need and level of familiarity. In this illustration, TM5 is associated with a low endurance need, with which the user has a great deal of familiarity.

The relationship between behavioral and topical models is somewhat circuitous. The creation of topical models grows out of the observation of a user's information search and use behavior. Once topical models have been identified and associated with particular behavioral models, the new behaviors exhibited by the user while engaged in activities associated with a particular topical model can be used to iteratively refine and update the associated behavioral

Kelly, D. & Belkin, N. J. (2002). A user modeling system for personalized interaction and tailored retrieval in interactive IR. In *Proceedings of Annual Conference of the American Society for Information Science and Technology (ASIST '02)*, Philadelphia, PA, 316-325.

model. Just as behavioral models are iteratively refined and updated over time, so too are topical models.

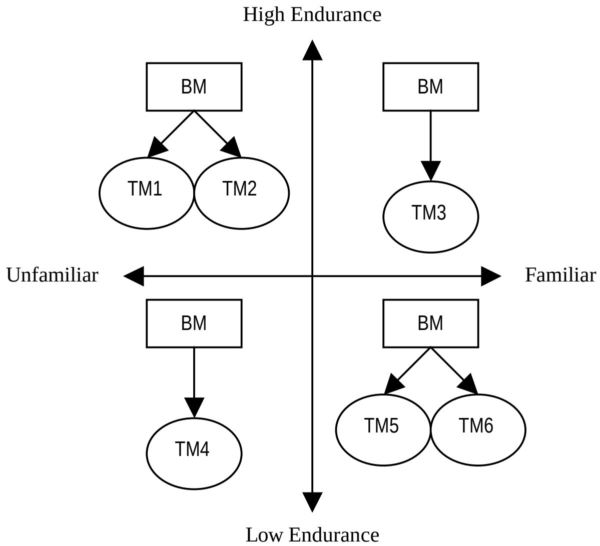

Figure 6. Relationship between personalized behavioral models (BM) and topical models (TM)

Topical models can be initialized and updated retroactively. Consider a user who enters a query that cannot be matched to a pre-existing topical model, but later performs some other activity that ties the interaction to a pre-existing topical model. For instance, the user might save a document in a folder that is associated with a topic model. In this case, the system recognizes at a later point during the search session that the user's information search and use activities are associated with a pre-existing topical model. The system can retroactively incorporate what it has learned about the user during the current search session into the appropriate topical model.

Over time, topical models may be refined, diminished or stabilized. The user's information search and use behaviors provide information on which to maintain and update topical models. If the frequency of use of a topical model for a user's information need diminishes, then the weight of that model as a retrieval aid diminishes. As the user's familiarity with a topic increases, as reflected in his/her information search and use behavior, so too will the types of information contained in the topical model.

Topical models can provide the system with a means of tracking previous information search and use activities and interactions with information objects. Accordingly, topical models should be informed by search histories of the user's previous information seeking episodes, as well as the particular query terms and documents that the user inputs and selects. Thus, topical models can be used by the system to understand something about what the user has learned through interacting with information objects across time. The user does not have to explicitly establish a context each time a search is initiated. Instead, topical models can function to establish a shared understanding between the system and user about what the user knows about a topic. Topical models attempt to capture the user's tacit knowledge about a topic through tracking and classifying the user's information search and use behaviors.

Topical models represent what the user "knows" about a topic. Because topical models are created over time, across multiple information seeking episodes, it is possible for the topical model to provide the system with a context in which to interpret the user's activities. Topical models can function to provide the system with a unique context of use for each user, by capturing the user's changing states of knowledge about a topic. Through the provision of unique language models for each user, topical models can be used to disambiguate term usage and select appropriate documents for retrieval. Finally, topical models can provide the system with evidence in which to customize the search interface based upon previous search interactions.

Conclusions

While UM offers the potential of personalized interaction and tailored retrieval, UM systems typically only consider a single, static information need of the user. UM systems do not acknowledge or account for changes in the user's states of knowledge. Furthermore, techniques for knowledge elicitation are cumbersome, intrusive, imprecise and incomplete. It is not enough to simply assemble a list of keywords that describe a user's interests. Instead, a UM system should track the cognitive progression of the user across search sessions in an attempt to understand how these search sessions are related, how term usage and needs vary and change over time and how the user's activities reflect an evolving state of knowledge.

If information search and use behaviors are to be used as implicit sources of information about the user, then these behaviors should be combined to form personalized behavioral models that model an individual's behavior and not the behaviors of a canonical group of users. We know from years of information seeking behavior research that people exhibit different searching behavior and that behavior is affected by a multitude of factors. Thus, we need a UM system that is personal both with respect to user behaviors and information needs. We should not build a UM system that relies on stereotypical user behavior any more than we should build a UM system that relies on stereotyping user's interests. If the goal is to provide personal interactions and tailored retrieval through implicit modeling techniques, then we need to create a UM system that considers both an individual's information seeking behaviors and information needs.

We have presented a framework for constructing user models for IR systems, based on observations of the user's behaviors over time. This framework has the potential advantage over other methods that have been used for user modeling in IR systems in that it depends upon implicit information, rather than on direct elicitation of information from the user. It also has the potential advantage of providing a means for developing and, more importantly, using long-term user models in a consistent way, which could personalize the IR interaction for each time each user engages in information seeking.

At the moment, this framework is only a proposal, which needs to be empirically investigated. Such investigation is currently under way, as a part of the MONGREL project being conducted jointly by Rutgers University and the University of Massachusetts, Amherst (MONGREL). We have initiated a longitudinal, naturalistic study of the information search and use behaviors of several users and will initiate a second study with additional users in the future. We hope that the results from these studies will indicate the effectiveness of our model for personalizing interaction and tailoring retrieval in interactive IR.

ACKNOWLEDGMENTS

The work was in part funded by NSF Grant #99-11942. Any opinions, findings and conclusions or recommendations expressed in this material are the authors and do not necessarily reflect those of the sponsor.

REFERENCES

Allen, R. B. (1997). Mental models and user models. In M. Helander, T K. Landauer, & P. Prablu (Eds.), Handbook of human-computer interaction (2nd ed., pp. 49-63). New York, NY: Elsevier.

Belkin, N. J. (1984). Cognitive models and information transfer. Social Science Information Studies, 4, 111-129.

Belkin, N. J. (2000). Helping people find what they don't know. Communications of the ACM, 43(8), 58-61.

Belkin, N. J., Brooks, H. M., & Daniels, P. J. (1987). Knowledge elicitation using discourse analysis. International Journal of Man-Machine Studies, 27, 127-144.

Belkin, N. J., Oddy, R. & Brooks, H. (1982). ASK for information retrieval: Part 1. Journal of Documentation 38(2), 61-71.

Billsus, D. & Pazzani, M. J. (1999). A personal newsagent that talks, learns and explains. Proceedings of the 3rd International Conference on Autonomous Agents (pp. 268-275). Seattle, WA.

Billsus, D., Pazzani, M. J., & Chen, J. (2000). A learning agent for wireless news access. Proceedings of Intelligent User Interfaces (pp. 33-36). New Orleans, LA: ACM Press.

Brajnik, G., Guida, G., & Tasso, C. (1987). User modeling in intelligent information retrieval. Information Processing & Management, 23(4), 305-320.

Croft, W. B., & Thompson, R. H. (1987). I^3R: A new approach to the design of document retrieval systems. Journal of the American Society for Information Science, 38, 389-404.

Daniels, P. J. (1987). Developing the User Modeling Function of an Intelligent Interface for Document Retrieval Systems. Ph.D. Thesis, The City University, London.

Ingwersen, P. (1982). Search procedures in the library analyzed from the cognitive point of view. Journal of Documentation, 38, 165-191.

Kamba, T., Sakagami, H. & Koseki, Y. (1997). ANATAGONOMY: A personalized newspaper on the World Wide Web. International Journal of Human-Computer Studies, 46, 789-803.

Kelly, D. & Belkin, N. J. (2001). Reading time, scrolling and interaction: Exploring implicit sources of user preferences for relevance feedback during interactive information retrieval. Proceedings of the 23rd Annual International Conference on Research and Development in Information Retrieval (ACM SIGIR '01)(pp. 408-409). New Orleans, LA. ACM Press.

Kelly, D. & Belkin, N. J. (2002). Reading time as an implicit measure of relevance in an information-seeking task. Manuscript submitted for publication.

Kelly, D. & Cool, C. (2002). The effects of topic familiarity on information search behaviors. In Proceedings of the Second ACM/IEEE Joint Conference on Digital Libraries (JCDL '02), Portland, OR. ACM Press.

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L.R. & Riedl, J. (1997). GroupLens: Applying collaborative filtering to usenet news. Communications of the ACM, 40(7), 77-87.

Lam, W., & Mostafa, J. (2001). Modeling user interest shift using a bayesian approach. Journal of the American Society for Information Science, 52(5), 416-429.

Lieberman, J. (1995). Letizia: An agent that assists web browsing. Proceedings of the International Joint Conference on Artificial Intelligence (pp. 924-929). CA: AAAI Press.

Lin, S. J. & Belkin, N. J. (2000). Modeling multiple information seeking episodes. Proceedings of the annual meeting of the American Society for Information Science, USA (pp. 133-146).

Mladenic, D. (1999). Machine learning used by Personal WebWatcher. Proceedings of the ACAI-99 Workshop on Machine Learning and Intelligent Agents. Chania, Crete.

MONGREL. Home page of the NSF-funded MONGREL project. http://www.scils.rutgers.edu/mongrel

Morita, M., & Shinoda, Y. (1994). Information filtering based on user behavior analysis and best match text retrieval. Proceedings of the 17th Annual International Conference on Research and Development in Information Retrieval (ACM SIGIR '94) (pp. 272-281). Dublin, Ireland. ACM Press.

Nichols, D. M. (1997). Implicit ratings and filtering. Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, 10-12.

Oard, D. W., & Kim, J. (1998). Implicit feedback for recommender systems. Proceedings of the AAAI Workshop on Recommender Systems (pp. 81-83).

Oard, D. W., & Kim, J. (2001). Modeling information content using observable behavior. Proceedings of the 64th annual meeting of the American Society for Information Science and Technology, USA. (pp. 38-45).

Oddy, R. N. (1977). Information retrieval through man-machine dialogue. Journal of Documentation, 33(1), 1-14.

Rich, E. (1983). Users are individuals: Individualizing user models. International Journal of Human-Computer Studies, 51, 323-338.

Rucker, J., & Polanco, J. (1997). Siteseer: Personalized navigation for the web. Communications of the ACM, 40(3), 73-75.

Saracevic, T., Spink, A., & Wu, M. M. (1997). Users and intermediaries in interactive information retrieval (IR): What are they talking about? User Modeling. Proceedings of the sixth international conference UM97 (pp. 43-54). NY: Springer.

Seo, Y. W., & Zhang, B. T. (2000). Learning user's preferences by analyzing wed-browsing behaviors. Proceedings of the ACM 4th International Conference on Autonomous Agents (pp. 381-387). ACM Press.

Spink, A. (1996). Multiple search session model of end-user behavior: An exploratory study. Journal of the American Society for Information Science, 47(8), 603-609.

Spink, A., Goodrum, A., Robins, D. & Wu, M. M. (1996). Search intermediaries elicitations during mediated online searching. Proceedings of the 19th Annual Conference on Research and Development in Information Retrieval (ACM SIGIR '96) (pp. 120-127). Zurich, Switzerland. ACM Press.

Spink, A., Griesdorf, H., & Bateman, J. (1999). A study of mediated successive searching during information seeking. Journal of Information Science, 25(6), 477-487.

Vickery, A., & Brooks, H. M. (1987). PLEXUS: The expert system for referral. Information Processing & Management, 23(2), 99-117.

Kelly, D**.** & Belkin, N. J. (2002). A user modeling system for personalized interaction and tailored retrieval in interactive IR. In *Proceedings of Annual Conference of the American Society for Information Science and Technology (ASIST '02)*, Philadelphia, PA, 316-325.