Chapter 3

Theoretical Framework

This chapter presents the theoretical framework for this dissertation. The first section presents a conceptual model of how a user modeling system incorporating implicit feedback as its primary mechanism for profile acquisition might function. This is only a basic model; specific technical details of how the system would work is left for future research, once empirical work supporting the general framework has been conducted. The presentation of this model is followed by a presentation of the research questions for this dissertation, which detail the specific aspects of the conceptual model that will be investigated. This is followed by a description of the study design. The conceptual model, research questions and study design are motivated by the problems with implicit feedback research presented in the previous chapter, including problems related to how behavior-based metrics used for implicit feedback have been collected, measured, interpreted and applied.

3.1 Conceptual Model

The conceptual model for this dissertation is based on the following desiderata for an ideal user modeling system for interactive information retrieval, criteria that have been developed based on an examination of previous research. A user modeling system for personalized interaction in information retrieval should:

- Acquire and update the user model automatically, without explicit assistance from the user.
- Account for information-seeking context.

• Personalize interactions to the individual.
• Track information needs and interactions over time.
• Represent multiple information needs, both short and long term.
• Allow for changes in information needs over time.

To achieve the goals above, two major classes of models are proposed: general behavioral and personalized. The general behavioral model describes how information-seeking behavior can be used to identify and track information needs. The characterizations found in this model are general to all users. The personalized model contains two types of models: personal behavioral and topical. The personalized model characterizes an individual user’s information-seeking behavior with regard to document preference and information-seeking context (personal behavioral), and the user’s information-seeking needs (topical). The characterizations found in this model are specific to the individual user. The personal behavioral models provide the system with formalizations about when it can infer document preference based on a user’s behaviors. Further, the personal behavioral models include some aspects of context, such as how familiar the user is with a topic and how long the user expects to be interested in information about the topic. Topical models include representations of the user’s information needs. It is proposed that topical models be inferred through characterizations defined by the personal behavioral models that describe the relationship between a user’s information-seeking behavior and document preference. Each type of model, the general behavioral and personalized, are described in more detail below. The discussion of the personalized model includes a discussion of both the personal

behavioral and topical models. Figure 3.1 displays the relationship between these various types of models.

![Diagram](image)

**Figure 3.1** A UM system for personalized interaction

### 3.1.1 General Behavioral Model

The General Behavioral Model (GBM) describes how implicit feedback can be used to identify and track a user’s information needs. The GBM contains characterizations of information-seeking behavior that can be used as starting points upon which to begin the construction of personalized user models. The relationships described by the GBM can function as hypotheses that are used by the system during the construction and maintenance of personalized user models to reason about an individual user’s behavior and the relationship of this behavior to the user’s document preference and information-seeking context. 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. 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 will be discussed in greater detail in the sections below as it relates to personalized models.

3.1.2 Personalized Models

Personalized models contain both personal behavioral and topical models, and 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 with 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-seeking behaviors, document preference and information-seeking context will be iteratively modified and adapted at the individual level as the system learns more about the user through observations that occur across time, and by the user providing some explicit feedback about their preferences.

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 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 equal weights. 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. 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 3.2. 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 information-seeking context. These are behaviors that User A exhibits with some regularity. For User B, the valid sources of implicit feedback for document preference and information-seeking context 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 information-seeking context. 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 information-seeking context. Figure 3.2 illustrates how personalized behavioral models might differ from user to user.

3.1.3 Accounting for Information-seeking Context

Because research on information-seeking has demonstrated that context can affect information-seeking behavior, it is important for the GBM to identify and account for contextual factors, such as topic familiarity and the length of a time a user expects to be interested in information on a topic (endurance), and to describe how this affects the information-seeking behaviors that the user exhibits. For instance, some behaviors may only be exhibited by a user in certain searching environments or with certain types of tasks or topics. Behavioral models for two example contextual factors, endurance of need and familiarity, is displayed in Figure 3.3. The behaviors represented in this Figure are the same as those represented in Figure 3.2.

---

Figure 3.3 Behavioral models for endurance of need and familiarity

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 contextual factors, such as familiarity and endurance of information need. One can imagine different values for various behaviors based upon these two dimensions. For instance, it might be the case that collections searched is a useful source of implicit feedback of user preference for an information need with high endurance; but not so for an information need with low endurance. Of course, some low endurance needs may eventually turn into high endurance needs and the behavioral model might 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 behavioral model for low endurance needs for that particular user can evolve. If the user continues to search for this need over time, then this user’s information-seeking

behaviors 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-seeking 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 display 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, display 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, display time for relevant and non-relevant documents may be different. In this situation, display 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 3.1 displays an example of the range of values that the GBM might contain for display 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-seeking behaviors about a given topic, the more familiar the topic becomes. The example in Table 3.1 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 display time for relevant and non-relevant documents. It is important to note that Table 3.1 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 3.1 illustrates how display time varies for relevant and non-relevant documents with respect to familiarity; as the user becomes more familiar with the topic, observed display times for both types of documents decreases.

Table 3.1

<table>
<thead>
<tr>
<th>Level of Familiarity</th>
<th>Relevant</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 (high)</td>
<td>12 seconds</td>
<td>5 seconds</td>
</tr>
<tr>
<td>2 (medium)</td>
<td>17 seconds</td>
<td>14 seconds</td>
</tr>
<tr>
<td>1 (low)</td>
<td>20 seconds</td>
<td>25 seconds</td>
</tr>
</tbody>
</table>

The range of values suggested by the GBM for display time can be used during the initial construction of the personalized behavioral models. Figure 3.4 demonstrates how this might occur using display 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.

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 3.2 displays User A’s behavioral model for low endurance needs, as well as some other models of display time that the system might have for User A. During the initial interactions, the system’s observations and interpretations of the user’s information-seeking 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 displaying those documents that he/she identifies as relevant and 9 seconds displaying 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 behavioral model for low endurance needs.

Table 3.2

Models of Display Time (DT) in Seconds, of Relevant (R) and Non-relevant (NR) Documents for User A

<table>
<thead>
<tr>
<th></th>
<th>Low Endurance Needs</th>
<th>High Endurance Needs</th>
<th>High Familiarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>NR</td>
<td>R</td>
</tr>
<tr>
<td>DT</td>
<td>11</td>
<td>9</td>
<td>20</td>
</tr>
</tbody>
</table>

Other examples of possible observed display times for User A and the models that they activate are displayed in Table 3.3. This table is used in conjunction with Table 3.2 and Figure 3.4. Observing display times of 11 seconds for relevant documents and 9 seconds for non-relevant documents would be matched to User A’s pre-existing behavioral model for low endurance needs, while observing display times of 20 seconds for relevant documents and 25 seconds for non-relevant documents would not be matched to any pre-existing 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 display times of 20 and 25 seconds would initiate the construction of this model based on the values found in the GBM.

Table 3.3

Examples of Display Time (DT) in Seconds, Observations and Initiated Action (BM: Behavioral Model)

<table>
<thead>
<tr>
<th>DT Observations</th>
<th>Matches BM?</th>
<th>Matches GBM?</th>
<th>System Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>NO</td>
<td></td>
<td>Initialize Low Endurance Model</td>
</tr>
<tr>
<td>NR</td>
<td>YES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td></td>
<td>Initialize High Familiarity Model</td>
</tr>
<tr>
<td>20</td>
<td>22</td>
<td>YES</td>
<td>Define Low Familiarity Model</td>
</tr>
<tr>
<td></td>
<td>NO</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

3.1.4 Topical Models

This section describes the final component of the proposed user modeling system, the component for topical modeling. 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 terms and 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-seeking 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-seeking behavior.

behavior. Thus, topical models are inferred based on the user’s information-seeking behavior. Behavioral models provide formalizations 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 topical models is based upon the user’s past information-seeking 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 the system is able to relate to a pre-existing topical model for that user, Topical Model 5 (TM5), which has some history of information-seeking interactions associated with it. In other words, previous information-seeking behavior has been associated with this model. Figure 3.5 is an extension of Figure 3.3. Figure 3.5 depicts the relationship between behavioral models for endurance of need, familiarity, and topical models. In this illustration, TM5 is associated with a low endurance need, with which the user has some familiarity.

Figure 3.5  Relationship between 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 particular topic model. In this case, the system recognizes at a later point during the search session that the user’s information-seeking 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-seeking behaviors provide information on which to maintain and update topical models. If the frequency of use of a topical model 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-seeking 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-seeking 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 by the particular query terms and documents that the user enters 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-seeking 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.

### 3.2 Research Questions

The success of the user modeling system described above rests on the assumption that there is a link between the behaviors a user exhibits during information-seeking, and that user’s preferences for documents. Empirical research, as discussed in Chapter 2, has found mixed results with regard to this assumption, although in general, most research is somewhat optimistic about the existence of this relationship. However, numerous problems have been identified with this research, including problems related to how behavior-based metrics have been collected, measured, interpreted and applied. Thus, this dissertation is concerned with several aspects of the user modeling system proposed in the above Section 3.2 that relate to these issues: understanding which behaviors can be used as implicit feedback, how these behaviors are affected by information-seeking context, and how these behaviors can be used to infer document preference. The purposes of this dissertation research are to: (1) collect and measure information-seeking behaviors using a valid, reliable method that optimizes ecological validity; (2) evaluate the predictive power of information-seeking behaviors as implicit sources of document preference; and (3) determine how information-seeking context affects the occurrences of these behaviors. This study is exploratory in nature and therefore, no specific research hypotheses are proposed. Instead, the following questions guide this research:

- What kinds of information about a user’s preference for a document can be gathered implicitly through monitoring the user’s information-seeking and use behavior and examining the results of that behavior over time?
- What is the relationship between a user’s information-seeking and use behaviors and their document preference?

---

• What is the relationship between a user’s information-seeking and use behaviors and the information-seeking context?

• How can a user’s information-seeking and use behaviors be used to predict the user’s document preferences?

In the following section, the study design is presented, along with details of how information-seeking and use behaviors, information-seeking context and document preference were defined in this study. The presentation includes those design decisions that were directly related to the goals and questions of this study; a more detailed discussion of the method and how particular variables were measured is presented in Chapter 4.

3.3 Study Design

This research observed and recorded subjects’ information-seeking activities and behaviors over an extended period of time, in a naturalistic setting; measured various aspects of the user’s information-seeking context; and obtained users’ preferences for the documents that they viewed. The purpose of this study was to collect, measure and relate aspects of the information-seeking context such as task and topic, to explicit behaviors such as display time, printing, saving, and bookmarking, and to evaluate the power of these aspects and behaviors in predicting document preference.

The methodological approach for this study was guided not only by the goals of this study, but also by the observed limitations of previous studies of implicit feedback. A longitudinal, naturalistic approach was selected to optimize ecological validity, by providing subjects with an opportunity to engage in multiple information-seeking

episodes with tasks and topics that were germane to their personal interests, in familiar environments. This approach was further selected so that changes in interests and behaviors over time could be measured and observed. A case study approach was used to focus on understanding the behaviors and preferences of individual users as opposed to a sample of users. This approach allowed for a detailed and comprehensive collection and examination of an individual’s behaviors and preferences. The selection of a time frame for this study was guided by practical purpose; it was determined that the study would last for an entire university semester (14 weeks) and that periods of measurement would take place at weekly intervals.

Because users often conduct information-seeking activities across disparate machines in distributed locations, laptop computers were provided to subjects. For subjects, the provision of laptop computers allowed for a locus of attention in which to conduct their activities, and acted as substantial incentive for granting access to the investigator to record and examine their activities. For the conduct of the study, the laptop computers provided consistency in implementing and maintaining the technical aspects of the study, including the installation and use of instruments for data collection.

### 3.3.1 Information-seeking Behaviors

Information-seeking behaviors were those behaviors that the subject naturally exhibited during online information-seeking activities. Client-side logging software was used to gather information on the occurrences of these behaviors. For this study, the following behaviors were measured and analyzed:

---

• Display Time. Display time was the length of time that a document was displayed in the subject’s active browser window.

• Retention. Retention behaviors referred to several actions that the subject might perform to keep a document for potential later viewing. Retention behaviors of interest to this study were printing, saving, and bookmarking.

These behaviors were selected for both practical and theoretical reasons. From a practical standpoint, display time and retention could be obtained relatively easily using commercially available client-side logging software. Cost prohibited designing and building a custom logger to collect additional behaviors. Technical issues also prevent the valid and reliable collection of some behaviors such as scrolling (Goecks & Shavlik, 2000). Furthermore, display time and retention were more easily interpretable than others; using a behavior like create would require the development of techniques for understanding and analyzing text. From a theoretical standpoint, display time has received a great deal of interest in the research community and the results regarding its potential as implicit evidence have been mixed (c.f. Claypool, et. al, 2001; Kelly and Belkin, 2001a; Morita & Shinoda, 1994; White, et. al, 2002). Retention was selected because it is often used as an evaluation metric for the potential of other behaviors as implicit evidence (Rafter and Smyth, 2001). This study sought to understand each of these behaviors in greater detail.

### 3.3.2 Information-seeking Context

Information-seeking context was defined in this study by a number of variables: task, topic, endurance, frequency, stage, persistence and familiarity. Endurance,
frequency and stage were considered to be attributes of task, while persistence and familiarity were considered to be attributes of topic. The selection of these variables was based on the literature presented in Section 3.2.1. Each of these variables will be discussed in more detail below.

Task was defined for this study as the goal of information-seeking behavior, and topic was defined as the specific subject within a task. One of the goals of the design of this study was to observe subjects in natural online information-seeking situations. Because of this, no pre-selected tasks or topics were assigned to subjects, and subjects were not presented with this definitions during the study. Instead, subjects were instructed to conduct their normal information-seeking activities. To provide structure to these activities and to understand more about what was trying to be accomplished, subjects were asked to think about their online information-seeking activities in terms of tasks and topics, to create labels for each task and topic, and to classify the pages that they viewed according to these tasks and topics. Subjects’ tasks and topics were elicited at the start of the study and at subsequent weekly meetings.

Subjects’ choice of tasks and topics was completely subjective. Indeed, anything that a subject believed was a task or topic was permitted. Subjects were provided with some example tasks and topics at the start of the study, but in no other way were they directed, influenced or biased in their choice of tasks and topics. The goal in doing so was to understand subjects’ personal interpretations of their tasks and topics, and subjects further elaborated on these interpretations during an Exit Interview. From a theoretical perspective, this methodological approach allowed for an exploration of user-specific

tasks and topics, techniques for their elicitation, and an understanding of their relationship to the subject’s information-seeking behaviors.

Along with subjects’ tasks and topics, various attributes of each task and topic – attributes that have been suggested by the literature (see Section 3.2.1) as comprising information-seeking context and influencing information-seeking behavior – were measured. For each task, subjects were asked to indicate the task endurance, frequency and stage; for each topic, subjects were asked to indicate the topic persistence and their familiarity with the topic. For each week of the study that the task or topic was active, subjects were required to reconsider and rate each of these attributes. Each attribute is defined below:

- Task endurance was the length of time that the subject expected to be working on a task.
- Frequency was how often the subject expected to conduct online information-seeking activities related to a task.
- Stage was the subject’s assessment of their progress in completing the task.
- Topic persistence was the length of time the subject expected to be interested in information about a topic.
- Familiarity was the subject’s current state of knowledge about a topic.

### 3.3.3 Document Preference

Subjects’ preference for a document was measured by how useful they believed the document to be in helping them to complete and/or understand the particular task and topic in which they classified the document. The decision to collect data about the

---

usefulness of a particular item as opposed to its relevance was based on the concept of utility as described by Cooper (1971). Cooper’s (1971) conception of utility allowed for a more user-centered, subjective assessment of relevance and diverged from a strict objective assessment. As stated in Chapter 2, this dissertation adopts the situational view of relevance as proposed by Saracevic (1996) and Wilson (1973).

In addition to usefulness, subjects were asked to indicate the navigational usefulness of the documents that they viewed. Navigational usefulness was added after interviews from a pilot study revealed that subjects were marking documents as useful because they “helped them get to a good page.” Because this study was concerned with unobtrusively identifying documents that subjects found useful from a content perspective, or those that could be used to model their interests, navigational usefulness was not of theoretical interest and was not considered during analysis. However, subjects were instructed to distinguish between these two types of usefulness. Navigational usefulness was described to subjects as, “A document is navigationally useful if it helps you get to another useful document.”

Subjects’ were asked to indicate their confidence with respect to the usefulness ratings that they assigned documents. Confidence was defined as how much the subject believed that their usefulness ratings were accurate.