Modeling Search Processes using Hidden States in Collaborative Exploratory Web Search

Zhen Yue
University of Pittsburgh
135 North Bellefield Ave.
Pittsburgh, PA, 15260
Zhy18@pitt.edu

Shuguang Han
University of Pittsburgh
135 North Bellefield Ave.
Pittsburgh, PA, 15260
shh69@pitt.edu

Daqing He
University of Pittsburgh
135 North Bellefield Ave.
Pittsburgh, PA, 15260
dah44@pitt.edu

ABSTRACT

Investigations of search processes that involve complex interactions, such as collaborative search processes, are important research topics. Previous approaches of directly applying individual search process models into collaborative settings have proven to be problematic. In this paper, we proposed an innovative approach to model collaborative search processes using Hidden Markov Model (HMM), which is an automatic technique for analyzing temporal sequential data. Obtained through a user study, the data used in this paper consist of two different tasks in both collaborative exploratory Web search and individual exploratory Web search conditions. Our results showed that the identified hidden patterns of search process through HMM are compatible with previous well-known models. In addition, HMM generates detailed information on the transitions of hidden patterns in search processes, which demonstrated to be useful for analyzing task differences, and for determining the correlation of search process with search performance. The findings can be used for evaluating collaborative search systems as well as providing guidance for the system design.

Author Keywords
Collaborative information behavior; exploratory search; Hidden Markov Model; information seeking process.

ACM Classification Keywords
H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces – Collaborative computing, Computer-supported cooperative work.

MOTIVATION AND BACKGROUND

A recent survey reported that the percentage of respondents engaged in collaborative Web search on a daily basis had increased from 0.9% in 2006 to 11% in 2012 [25] . As the needs of collaborative exploratory search continue to gain more and more attentions, researchers have developed many new systems and interfaces to support collaborative information seeking and retrieval [2, 9, 15, 26, 33]. Studies that seek to investigate and model search processes had played important roles in the development of Web search interfaces for individuals. For example, Bates [3] incorporated an information seeking model into a real search interface designing, and Hearst [16] emphasized the necessity to understand human information search processes in designing successful search systems. We assert that the design of a collaborative search system can benefit from a study of collaborative search processes. Previous researches [18, 32] found that it is inappropriate to directly apply individual information seeking models in collaborative search. This is because collaborative search involves more complex interactions--not only the interactions between users and the system but also the communication among users. We need new approaches that can tackle the complexity of interactions.

In individual search, researchers had employed two major approaches to investigate the information search process. One approach focuses on qualitative constructs, such as stages and context in the search process. Kuhlthau's model [20] and Marchionini's [22] model adopted such an approach. The other approach tries to derive search patterns through the analysis of logged user behaviors. These studies are based on units of search behaviors, such as actions [7, 8, 20], search tactics [42] or search strategies [4, 22], which are labeled either by human raters [22, 42] or automatic methods [7, 8, 20]. The manually labeling method is more controllable but requires huge human efforts and is difficult for scaling to large datasets. Automatic methods [7, 8, 20] need fewer human interventions, but these methods make it difficult to connect the outputs with existing theoretical models.

To investigate the collaborative search process, both [18] and [32] attempted to map the individual information search model into a collaborative Web search; both papers found this approach to be inappropriate. Although not specifically targeting a collaborative Web search, Evans and Chi's social search model [12] benefits the study of the collaborative search process. The model was built upon survey data rather than logged user data. Considering the complexity of interactions involved in the collaborative search, studying collaborative search processes through logged user behaviors remains a problem.

CSCW'14, February 15 - 19 2014, Baltimore, MD, USA

http://dx.doi.org/10.1145/2531602.2531658

In this paper, we focus on automatic methods for analyzing the collaborative information search process using logged user behavior. Previous studies on investigating the search process using logged data either focuses on the observed action level [7, 8, 20] or on the manually coded search tactics/strategies level [22, 42]. The connections between the observed actions and search tactics/strategies are missing. To capture such connections, we took a two-level view of the basic unit in the search process in this paper. The first level is observable actions, which are the manifestation of users' behavior. The second level is unobservable search tactics or strategies. We assumed that users can move between different search tactics. In each tactic, users have a list of choices (each choice is an action), and each choice has its probability of being adopted by the user. The observed actions represent the users' adopted choices. Fuhr's theoretical framework for interactive retrieval [13], which assumes that a user moves between situations in interactive information retrieval, supports this assumption.

In order to model the temporal sequential user behaviors and simultaneously leverage these two levels, we proposed using the Hidden Markov Model (HMM) [21]. Because search tactics usually represent the users' internal choices and are difficult to observe, it is reasonable to model the search tactics as the hidden variables. Moreover, unlike the Markov models that have been employed [6, 38] to analyze individual search processes, HMM assumes a Markov chain on the unobserved hidden tactics rather than the observed user actions. This approach helps to remove the over-simplistic assumption in Markov models: user's future action depends only upon the present user action, not on the other actions before the present action. One additional benefit of HMM is that instead of manually labeling user actions; HMM is an automatic method that easily applies to large datasets.

Although HMM is a well-established model, the majority of its applications focus on predicting future events, such as the weather and stock prices. To the best of our knowledge, the use of HMM in analyzing search processes is rare; we know of no previous works that applied HMM in collaborative Web search. Therefore, little available information explains how to categorize the observable user actions, how to choose appropriate parameters and how to make sense and interpret the outputs. This leads us to focus on investigating the following three research questions:

RQ1. How to apply HMM in the analysis of information search processes, especially in collaborative Web search?

RQ2: How to interpret the outputs of HMM? How to determine whether HMM is a valid method for analyzing collaborative exploratory Web search process?

RQ3: What are the potential applications of the HMM outputs? Can these outputs be used to analyze the connections between the patterns of search processes and the search task difference or the search performance?

RELATED WORK

The literature related to our study can be divided into three areas: collaborative Web search and systems, individual search processes, and collaborative search processes.

Golovchinsky et al. [14] classified the collaboration in Web search using three dimensions – location, concurrency, and intent. Morris [24] surveyed 204 information workers about when they collaboratively used Web search tools and on what tasks they usually collaborated with others. Evans and Chi [12] also investigated 150 people using Mechanical Turk on collaborative and social search strategies. Their study revealed that collaborative Web search is a surprisingly common activity. A most recent survey [25] suggested that the increased prevalence of collaborative Web search is a result of the significant change of the technology landscape, particularly the rise of social networking sites and the growing usage of smartphones. The survey also suggested that users expressed frustration regarding their lack of awareness of collaborators' activities; as a result, redundant work emerges as the primary concern in current collaborative Web search practice. Shah and Marchionini [33] presented a user study to examine the awareness in collaborative information-seeking. They showed that support for group awareness is more effective than an awareness of personal actions and history. Morris [27] conducted a study to investigate the sensemaking in collaborative Web search. Her study pointed out that sensemaking is an important component in collaborative search processes, and compared several different system features for supporting collaborative sensemaking.

In individual search, researchers explored many theoretical models to investigate the search processes of single users. There are several well-known search process models, such as Kuhlthau's [20] and Ellis' [10] model. Marchionini [22] proposed an information search process model of eight stages with possible transitions between each stage. This model highlights the likelihood of one stage transitioning into another stage in three types: most likely transitions, high-probability transitions, and low-transition probabilities. Other studies look into the search patterns through the analysis of user-system interactions. Holscher and Strube [17] compared action sequences between Internet experts and "newbies". Chen and Cooper [6] used a stochastic model to examine usage patterns in a Web-based library catalog. These studies examined search processes through the actions in the log without recognizing user intentions behind those actions. Xie and Joo [38] raised the importance of investigating transitions of search tactics as a means of examining search processes. Using a predefined scheme of search tactics, they first manually coded the transaction logs. Then, they developed a five-order Markov chain to find the common search tactics in the behavior sequence of the user. Elsweiler et al. [11] proposed using the Markov model to investigate search processes in Personal Information Management

(PIM). Most of the above-mentioned works use the Markov model on the observed action level, which made the over-simplistic Markov assumption.

One way to overcome the inappropriate assumption is to model the action sequence on the unobserved level. For example, Fuhr [13] extended the classical Probability Ranking Principle (PRP) into Interactive Information Retrieval (IIR) contexts. Fuhr's model assumes that users move between situations. In each situation, users have a list of choices, and each choice (e.g. a user action) has its cost, acceptance probability, and benefit. However, this model only provided a theoretical framework without detailing how these parameters can be estimated. Later, Tran and Fuhr [36] combined both user click logs and eye-tracking data to model the search process as a Markov chain. However, the authors did not make a clear distinction between the situations and the actions. The Markov model is built on four pre-defined Areas of Interests (AOIs).

The Hidden Markov Model (HMM) is a well-established model with mature techniques for parameter estimation. The information retrieval area has used HMM. For example, HMM has been adopted in inferring users' search contexts [5] and predicting next user behavior [37]. However, the goal of both of these papers is to improve the performance information retrieval algorithm rather than investigating and interpreting the search process.

In collaborative search, Evans and Chi [8] proposed a social search model consists of before, during and after search stages and suggested the information exchange is valuable in the whole process. The model provides a holistic view of search process while a fine-grained model is still needed to provide more details. Several studies attempted to expand Kuhlthau's model in collaborative setting. Hyldegard [18] explored Kuhlthau's model in a group educational setting based on a qualitative preliminary case study. She found that collaborative search process cannot be modeled the same way as individual search process. She suggested that the Kuhlthau's model should be extended to incorporate the impact of social and contextual factors in relation to collaborative information seeking process. Shah and Gonzalez-lbanez [32] also attempted to map Kuhlthau's model to collaborative information search process. Similar to Hyldgard, they also declared that social elements were missing when applying the Kuhlthau's model in a collaborative setting.

Based on the above literature review, we can see that the investigation on collaborative search processes is limited and problematic. Markov chain on observed actions had been used for studies in individual search settings, but the oversimplified Markov property may be inappropriate for analyzing search processes that involve complex interactions, such as collaborative search processes. While HMM had been proved to be a more powerful approach than Markov chain, there was no study investigating the application of HMM in studying search processes.

EXPERIMENT DESIGN

CollabSearch: a Collaborative Search System

CollabSearch[1] is a web search system for either a single user or a group of users. As with other collaborative search systems, CollabSearch has both search and collaboration features. As shown in Figure 1, the left side of the system's interface offers a space for chatting. Team users can communicate with each other by sending instant messages. The chat feature is turned off and the chat box is hidden in individual search mode.

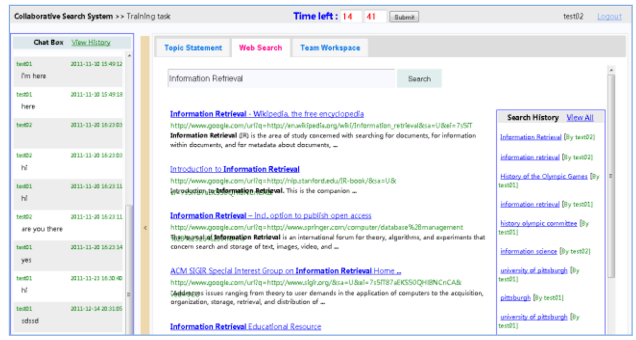

Figure 1: The web search frame

The main interface of CollabSearch contains three frames: topic statement, search and team workspace. The topic statement frame shows the current task description. In the navigation bar, they can see the time remaining for the current task. The search frame connects a user's query to the Google search engine, displaying the Google search results. This frame also reveals users' search history as well as that of their teammates. Users examine search results in the returned list for relevant information, and can save a whole web page or a snippet of the page.

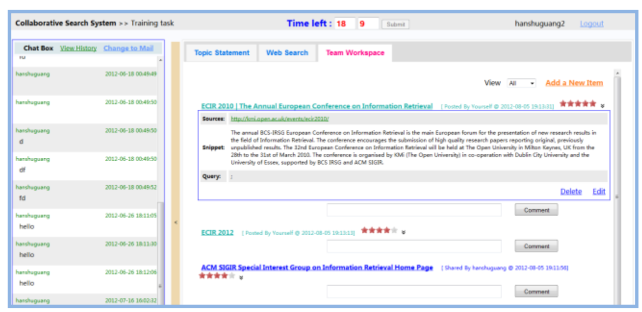

Figure 2: The workspace frame

All of the saved web pages and snippets, collected by the users in the same team, are stored in the team workspace frame. A notice is displayed at the top when new items have been saved to the team workspace. Users can click to view more details about an item in the workspace. Users can also decide whether a particular item is visible to other team members or not. The items saved by the user and the teammates are displayed in different colors and the user can choose to view a particular sets of items by using the filter.

[1] http:// crystal.exp.sis.pitt.edu:8080/CollaborativeSearch/

Experiment Conditions

We adopted a mix-method experiment with one within-subject factor: search task; and one between-subject factor: search condition, which refers to two search modes:

Collaborative search condition (COL). In this condition, two participants formed a team, and they worked on the same task simultaneously. As we were trying to simulate remotely-located collaboration, the team members could only communicate with each other by sending instant messages or reading each other's search histories. The collected results are shared in the team workspace, but no face-to-face communication was allowed.

Individual search condition (IND). This condition was devised as a baseline. In this condition, participants work on the exploratory search tasks individually.

Search Tasks

Two exploratory web search tasks were used in this study. Both of them had been used in other collaborative web search studies [23, 19]. One task is related to academic work, asking participants to collect information for a report on the effect of social networking service and software. The other one is the project-based leisure task [35], asking participants to collect information for planning a trip to Helsinki. Morris [24] noted that academic work and travel planning are the two common collaborative search tasks. Therefore, both of them are appropriate to study collaborative Web search.

The reason that we chose these two tasks is that they represent two different types of exploratory Web search tasks. First, the academic task (T1) is a recall-oriented information-gathering task, whereas the leisure task (T2) is a utility-based decision-making task. Second, the relevance criteria in these two tasks are different according to Saracevic's relevance theory [30]. Topical relevance is probably the most important criterion in T1 because the whole task is more objective in relevance judgments, whereas T2 involves users' subjective judgments and even personal preferences, so the relevance criteria contain many subjective and personal flavors.

Experiment Procedure

Each team in COL worked on both tasks. The order of the two tasks was rotated to reduce the learning and fatigue effect. During the experiment, participants were introduced to the study and the system, and completed an entry questionnaire to establish their search background. Then, participants worked on a training task for 15 minutes to become familiar with the system. They went on T1 or T2, depending on the task order assigned to each team. They had 30 minutes for each task. At the end of each task, participant completed a post-search questionnaire about their satisfaction on the performance and the cognitive load of search experiences. The experiment procedure in IND is identical to the COL condition except that only one participant undertaking the entire process.

Participants

Fifthy-four participants including 26 females and 28 males were recruited from the University of Pittsburgh. Among all of the participants, 36 signed up in pairs and thus formed 18 teams, which were assigned to the COL condition. The remaining 18 individual participants were assigned to the IND condition. All participants are students and they use computers and conduct Web searches on daily basis. Twenty-four participants are graduate students whereas the other 30 are undergraduates. According to a question asking them to rate their search experiences from 1-7 (with 1 as the least experienced and 7 as the most experienced), the response ranged from 4 to 7; thus most of our participants were experienced searchers.

Data Analysis Method

Since our study contains both within-subject and between-subject factors, and data maybe correlated, we adopted generalized estimating equation (GEE) to fit the model to the data, and analyzed the relationships between the independent variables and dependent variables, as well as the correlations. $p \leq .05$ was used to test any significant differences. GEE is a flexible statistical tool which deals with repeated measure and non-normal distributed data.

MODELING SEARCH PROCESS USING HMM

Categorizing user interactions

Before using HMM, we preprocessed user actions by classifying them into meaningful categories. Belkin et al. [4] classified user interactions using four dimensions: method of searching, mode of retrieval, goal of retrieval, and resource considered. The combination of dimensions defines multiple user interactions. In his model, each dimension was presented as binary values. Later Kim [19] expanded some dimensions with more than two values and removed some dimensions and values that do not apply to Web search environment. Xie [38] used two dimensions—methods and resources to classify user interactions, and she defined 8 values for methods and 6 values for resources.

For team users, the interactions are more complex than individual users. For example, when team users use CollabSearch for a collaborative Web search, the items saved in the workspace are both the ones saved by user him or herself and the partner. Therefore, it is useful to distinguish the interactions in which the user clicks the workspace to check on his/her own saved items from the interactions where the user checks on the partners' items.

Inspired by those ideas of classifying interactions using dimensions, we employed the following three dimensions to categorize user interactions in collaborative Web search: method, object and source, as shown in Figure 3. Some of the values of the method dimension, like search, scan, select and capture were also used in Kim's model [19]. However, we added another value communicate, which is unique in collaborative search. The values of the object dimension include all the possible objects that may exist in the collaborative Web search process, including query, topic statement, single item in search result, chat messages, list of saved items, and single saved items. The source dimension is unique in collaborative search context because it is important to distinguish the source of a particular object in collaborative process, whether it is from the user him/herself or from the partner or it is a shared/mix object.

Method	Object	Source
Search	Query	Self
Scan	Topic statement	Partner
Select	Item in search result	Shared/ mix
Capture	Chat messages	
Communicate	List of saved items	
	Single saved item	

Figure 3: Three dimensions for classifying user interactions

Using the combination of the following three dimensions, a set of observable user interactions is defined (see Table 1). Search-query-self (Q) represents the user issuing a query. In terms of using the workspace, users can scan the whole workspace without clicking on any particular item; this kind of interaction is the scan-list of saved item-mix (Wm). If the user clicks on a particular item, depending on whether the item is saved by the user or the partner, the interaction can be select-single saved-item-self (Ws) or select –single saved item- partner (Wp). All the possible combinations of the three dimensions that can be observed in the CollabSearch system are listed in Table 1.

Interactions	Description
Search – query – self (Q)	A user issues a query
Select- item-self (V)	A user clicks on a result in the returned result list
Capture-item-self (S)	A user saves a snippet or bookmarks a webpage
Scan-list of saved item – mixed (Wm)	A user checks the workspace without clicking on any particular item.
Select – single saved item –self (Ws)	A user clicks on an item in the workspace saved by him/herself
Select – single saved item – partner (Wp)	A user clicks on an item in the workspace saved by the partner
Scan-topic -shared (T)	A user clicks on the topic statement for view
Communicate-messages-self (Cs)	A user sends a message to the other user
Communicate-message-partner (Cp)	A user receives a message from the other user

Table 1: User interaction categorization

There are other possible combinations like select – item – partner, which means the user can click on an item in the partner's search result. However, this interaction is not supported in CollabSearch because team users do not share the screen and they cannot see the process when the partner issues a query and gets a returned results list. We acknowledge that clicking/viewing the topic statement is an artificial action that only exists because of the experiment setting. However, no matter whether it is because of the vagueness or the evolving nature of the exploratory information needs, users in exploratory search often have to consult their understanding of the search task to assess their progress in search. This consultation usually happens in users' mind without many external signals. Therefore, our topic statement action, though it sounds artificial, actually resembles a step that does exist in exploratory searches. We use the activities recorded in search logs to identify clicking/viewing topic statement.

Modeling Search Process

In this study, we introduced the Hidden Markov Model (HMM) method to model search tactics and search actions simultaneously. The model is described in Figure 4. We have a sequence of user actions from A_1 to A_M, and each action is one of those predefined nine actions: {Q, V, S, Wm, Ws, Wp, T, Cs, Cp}. HMM assumes that we also have a sequence of hidden states, from H_1 to H_M, and each action is generated by a corresponding hidden state, but different actions can be generated by the same hidden state with different probabilities. In this case, each action is corresponding to only one hidden state, and the hidden state sequence forms a Markov Chain.

A HMM model has several parameters: the number of hidden states N, the start probability of each states $\boldsymbol{\pi}$, the transition probabilities among any two hidden states $\boldsymbol{A_{ij}}$ and the emission probability from each state to each action $\boldsymbol{b_{ij}}$. By only defining the N and $\boldsymbol{\pi}$, a Baum-Welch algorithm [21] can be used to estimate the emission and transition probabilities.

Model Selection

It is still an open issue for determining the number of hidden states. Determining number of hidden states **N** is a model selection problem in learning the Hidden Markov Model. A complex model with large number of states will help to increase the sequence likelihood because there are more parameters that can be used to describe the model more precisely. But it has a high risk of causing over-fitting. A simple model is less likely to over-fit on the given dataset, but it may not be able to uncover the natural feature of datasets. In model selection, the information criterion such as the Akaike information criterion (AIC) or its variants [1] and Bayesian information criterion (BIC) [23] can be used to determine the optimal number of states. In this paper, we used BIC because it considers sample size.

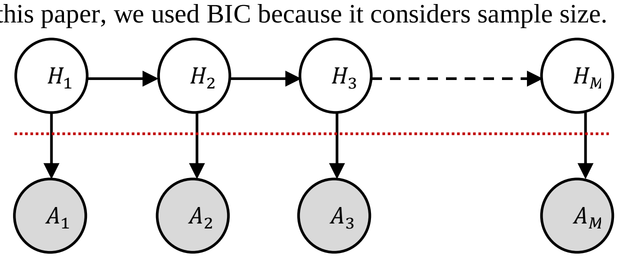

Figure 4: A Hidden Markov Model for Search Tactics

Suppose the number of parameters in HMM is p, and the number of total samples is s. The BIC is defined in

Formula (1), in which L denotes the log-likelihood of all samples. p can be computed using $(N-1)+(N-1)\times(N-1)+N\times(M-1)$, considering the summation of all probabilities is 1. M denotes the number of action types. A large log-likelihood with less parameters is preferred.

$$BIC = -2\times\log(L)+\log(s)\times p \qquad \text{Eq. (1)}$$

VALIDATION OF HMM

Validating HMM in Individual Search

In this section, we validated the application of HMM in the individual search process. Figure 5 plots the BIC values against the number of hidden states in the IND condition. We can see that BIC has the optimal value when the number of hidden states is set to 4.

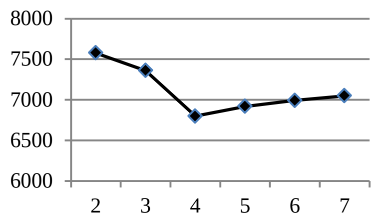

Figure 5: BIC Evaluation of HMM parameters in IND

Hidden states of HMM are represented by the emission probability distribution over observable user actions. The results of emission probability distribution in IND are shown in Table 2, in which we removed the probabilities that are smaller than 0.05 for better visualizing each hidden state. The first hidden state has a very high probability (0.99) of generating the interaction Q (defined in Table 1). Therefore, we named it HQ. Using the same naming criteria, we defined the second and third hidden states as HV and HS, respectively. It is clear that these three hidden states are directly related to search. The fourth hidden state is the most interesting one; it has a 0.57 probability of generating Ws and a 0.42 probability of generating T. We think that this hidden state is related to sensemaking, which is the process of bridging a knowledge gap that prevents the user from accomplishing the task [8]. In the exploratory search, participants may lack knowledge about the information problem, result space or the needed vocabulary for search [28]. In this hidden state, the participants were trying to evaluate the current search stage and define the current search problem in order to advance the search. Therefore, we named it the hidden state of defining search problem (HD).

	Q	V	S	Ws	T
HQ	0.99				
HV		0.91			
HS			0.96		
HD				0.57	0.42

Table 2: Hidden States and Emission Probability in IND

In HMM, each observed action corresponds to a hidden state in Table 2. To compare the differences between two tasks (this is for a later section), we computed the mean transition probabilities from each hidden state to any of the four (including itself) hidden states across all the participants. The transition probabilities are visualized in Figure 6 (probabilities lower than 0.05 are omitted). We can see a pattern of high transition probabilities on HQ→HV→HS, which represents a typical search pattern – a query is issued and results are viewed and saved if they are relevant. After saving an item, the participant may continue view another item HS→HV, issue another query HS →HQ, or transitioned to the sensemaking states before issuing another query HS→HD→HQ.

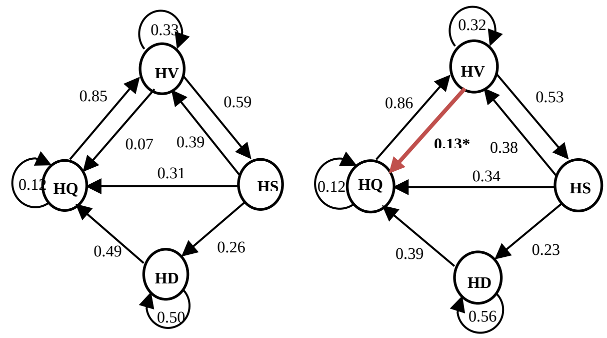

Figure 6: Comparison of transition probabilities of hidden states in IND for two tasks (left: T1, right: T2; a red arrow indicates significant difference: *p<0.05)

To validate HMM, we compared its output with Marchionini's information search process (ISP) model [22], which is a well-established model in the information seeking field. We found that the transitions among hidden states were very similar to the transitions among sub-processes in the ISP model. The default transitions in the ISP model can be mapped into HD→HQ→HV→HS→HD (Table 3), which is also the pattern of the highest transition probabilities in Figure 6. The ISP model also described the high and low transition probabilities among different sub-processes. For example, "extract information" (HS) had a high probability of transitioning to "examining results" (HV) and "formulate query" (HQ). Those transitions were also represented in the HMM output, with more details on the probabilities. Another model that can be used to validate the HMM result is the model of sensemaking loop [29]. The sensemaking loop is also reflected in the HMM output (transitions between sensemaking-related hidden states and search-related hidden states) with more details on the transition probabilities.

Sub-processes in the ISP model	HMM
Define Problem	HD
Select Source, Formulate Query, Execute Query	HQ
Examine Results	HV
Extract Information	HS
Reflect/Iterate/Stop	HD

Table 3: Mapping from sub-process in Marchionini's ISP model to the hidden states

Validating HMM in Collaborative Search

In this section, we validated the application of HMM in the collaborative search process. In Figure 7, we can see that the BIC has the optimal value when the number of hidden states in the COL condition is set to 6.

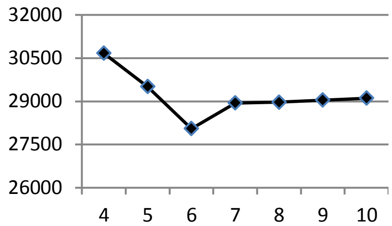

Figure 7: BIC Evaluation of HMM parameters in COL

The emission probabilities of each hidden state over observable interactions are shown in Table 4. The first three hidden states have high probabilities of generating Q, V and S respectively, which are similar to the first three hidden states in the IND condition. Therefore, we assigned them with the same names as in IND. These hidden states are directly related to the search while the rest three are related sensemaking. HD has a 0.36 probability of generating Wm, a 0.36 probability of generating T and a 0.21 probability of generating Cs. We think that this is a sensemaking hidden state in which the participants define the current search problem. Besides looking at the information in workspace and topic statement, the participants may also communicate with their partner to discuss the current search problem. The remaining two hidden states HW and HC are related to the communication between team members during the sensemaking stage. HW represents the state of checking the saved item details in the workspace whereas HC represents the continuous communication between team members.

	Q	V	S	Wm	Ws	Wp	T	Cs	Cp
HQ	0.82							0.13	
HV		0.87							0.1
HS			0.88						
HD				0.36			0.36	0.21	
HW					0.37	0.44			0.12
HC								0.44	0.47

Table 4: Hidden States and Emission Probability in COL

The transition probabilities were visualized for each search task respectively in Figure 8 (probabilities lower than 0.05 are omitted). Our results showed that sensemaking in the collaborative exploratory search (HD+HW+HC) were more important and complex than that in the individual search (HD only). We also recognized different types of sensemaking tactics, such as chat-centric sensemaking (HC) and workspace-centric sensemaking (HW). These results are consistent with the sensemaking types mentioned in [27]. In addition, our HMM outputs were consistent with Evans and Chi's model [12]. Their model consists of three phases: before search, during search and after search. In the before-search stage, users mainly focus on gathering requirements, which can be mapped to the HD in the HMM output. The during-search stage for informational tasks highlights the fact that "the foraging and sensemaking loops are tightly coupled," which is also reflected in the HMM output. The sharing of information with others in the after-search stage is also reflected in the HMM output with more details on what follows the sharing. The HMM output shows that after sharing, users could explicitly communicate about it or continue to the next round of defining a search problem.

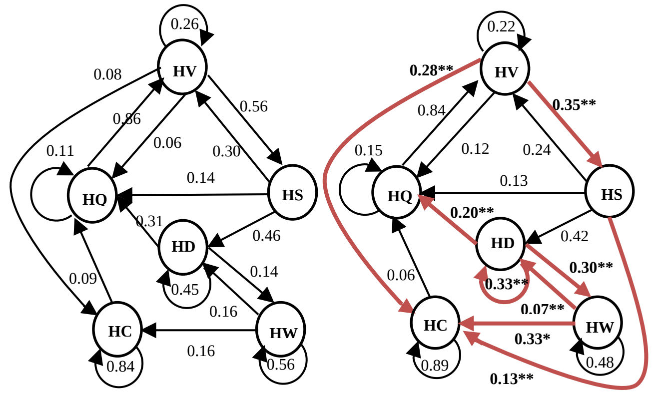

Figure 8: Comparison of Transition Probabilities of Hidden states in COL for two tasks (left: T1, right: T2; red arrows indicate significant difference: *p<0.05, **p<0.01)

The consistency of our HMM results with that of previous search process models in both individual search and collaborative search demonstrate the validity of HMM. The HMM results not only reveal the patterns found by previous models, but also provided more detailed information than previous model such as probabilities of transitions among different hidden tactics, which can be utilized to better understand users' search behavior.

APPLICATION OF HMM

The benefit of HMM is that it provides detailed information on the transition probabilities among different states in the search process. In this section, we showed the applications of such information through two cases.

Case I: Comparing Task Difference

Evans and Chi's model [12] built different search processes for three different types of user needs: informational, navigational and transactional. However, their model cannot distinguish two tasks within the same category. In this study, we showed that the HMM outputs can be used to reveal task differences.

The task comparison of individual search and collaborative search using HMM outputs are visualized in Figure 6 and Figure 8, respectively. In individual search, the patterns in the two tasks are very similar except HV→HQ, which we found it is significantly lower in T1 than in T2 (Mean diff = -.06, SE=.02, p=.012). In T2, participants were more likely to issuing another query after viewing a result, which might because the item was irrelevant. This may reflect the differences of task requirements. T2 is a decision making task, the participants are more concerned about the integration of information rather than collecting information. Therefore, they may be more selective on what information to save.

In the collaborative search, we can see that there are several significant differences between the two tasks, especially for the transitions between search related hidden states and sense-making related hidden states. HV→HC is significantly lower in T1 than in T2 (Mean diff = -.20, SE=.03, p<.001), so does HS→HC (Mean diff = -.10, SE=.03, p=.001). These differences suggest that when working on decision-making task T2, the participants were more likely to communicate with each other after they viewed or saved an item. Also, HW→HC is significantly lower in T1 than in T2 (Mean diff = -.17, SE=.07, p=.013). After viewing an item saved in the workspace, the participants were more likely to discuss what they think about the saved item in decision-making task T2. The transitions from HD are also different in the two tasks. HD→HQ is significantly higher in T1 than in T2 (Mean diff =.11, SE=.03, p<.001) whereas HD→HW is significantly lower in T1 than in T2 (Mean diff = -.16, SE=.06, p=.003). These results may indicate that in the information-gathering task such as T1, the participants preferred to have an overview of the workspace when they need to make sense of the current search problem whereas the users in decision-making task such as T2 preferred to look at the details of each saved items.

In both of the individual search and the collaborative search process, HMM recognized two types of hidden tactics, *i.e.* the search-related tactics and the sensemaking-related tactics. The search-related tactics remain similar in both search conditions, but sensemaking-related tactics are more complex in the collaborative search than that in the individual search. In terms of task difference, the collaborative search process exhibits more sensitivity on task differences. The ability of HMM on detecting task differences can be used for intelligent system design. When certain task type is detected, the system can provide support that is more suitable for the task.

Case II: Correlation between the Search Process and the Search Outcome

The ultimate goal of studying search process is to locate the core factors that influence search outcomes and provide better support for those factors. In this section, we are interested in locating those factors in collaborative web search process from real user behaviors, particularly how sense-making tactics influence the overall search performance.

Since we used the open Web as search collection, there is no ground-truth to calculate traditional precision and recall. In order to measure precision and recall, we defined the relevant Webpages (Ground-Truth data denoted by G) as those were saved by at least two teams. Then, for each participant u, the precision (P) and recall (R) were computed using Equation (2) and Equation (3), in which $S(u)$ are a set of webpages saved by participant u.

$$P(u) = \frac{|G \cap S(u)|}{|S(u)|} \qquad \text{Eq. (2)}$$

$$R(u) = \frac{|G \cap S(u)|}{|G|} \qquad \text{Eq. (3)}$$

An advantage of collaborative search is that users are able to discover hard to find webpages [31]. Likelihood of discovery has been used to measure how hard a Webpage can be found [31], which is inversely proportional to the number of teams visited the Webpage. However, it doesn't consider the quality of a webpage. And we defined a weighted likelihood of discovery (WLD) for each Webpage w by multiplying the likelihood of discovery of a Webpage by the quality of the Webpage, which is measured by the number of teams saved divided by number of teams visited. WLD is calculated using Equation (4), in which $S'(w)$ is the total number of teams saved w and $V'(w)$ is the total number of teams visited w. Since WebPages saved by at least two teams are considered as relevant, $S'(w) - 1$ and $V'(w) - 1$ is used.

$$WLD(u) = \sum\nolimits_{w \in S(u)} \frac{|S'(w) - 1|}{|V'(w) - 1|} \frac{1}{|V'(w) - 1|} \qquad \text{Eq. (4)}$$

User satisfaction (Sat) is a subjective measurement which reflects the participants' perception of the search outcome [34]. In this study, we measure the participants' satisfaction based on their responses to a question asking them to evaluate their satisfaction on the overall performance using 7-point Likert scale. A higher score indicates a better satisfaction. Cognitive load (Cog) can measure how hard one have to work for solving a complex problem, which can be used as a subjective measurement to evaluate the participants' perception towards their search experience. We used the instrument which has been used in another collaborative Web search study [31] to measure participants' cognitive load. A high cognitive load indicates a negative user perception to the search experience.

In this study, we are particularly interested in how sense-making are related to search performance. Therefore, we examined the correlation of the following transitions with the search performance, including HS→HD, HS→HC, HV→HC, HD→HQ and HC→HQ. These transitions are chosen based on two criteria: 1) the probability in either task is higher than 0.05; 2) it represents a transition between search and sense-making. We also considered HW→HC because it represents the transition between two different types of communication: from implicit communication to explicit communication.

	HS→HD	HD→HQ	HW→HC	HC→HQ
P	-	-	-	-
R	↓(p<0.01)	↑(p<0.01)	-	-
WLD	↓(p=0.01)	↑(p<0.01)	-	↑(p=0.06)
Sat	-	↑(p<0.01)		↓(p<0.01)
Cog		↓(p<0.01)	↑(p<0.01)	

Table 6: Correlation of Search Process and Search Outcome: ↓ denoting negative and ↑ denoting positive correlation

The results are shown in Table 6. ↑ means the transition is positively related to the performance and user perception while ↓ represents a negative relationship. We didn't find any significant differences on HV→HC and HS→HC, thus they are not shown in the table. We found that the transition from sense-making to search (HD→HQ) is positively related to performance and user perception. However, the transition from search to sense-making (HS→HD) is negatively related to performance and user perception. This might be caused by the fact that when participants are facing knowledge gap during exploratory search, they need to transit to sense-making states; and when the problems were solved, they transit back to the search states. Another interest finding is about HC→HQ. It's positively related to WLD but negatively related to Sat. It might suggest that explicit communication helped the participants to improve the search performance. However, explicit communication maybe triggered by a problem in search; which makes the participants feel less satisfied with their performance. In addition, we found that HW→HC is positively related to Cog, which means that a transition from implicit communication to explicit communication increases the participants' cognitive load. This might indicate that team members had something to negotiate through explicit communication, which increases the cognitive load.

Future Applications

We believe that there are still many applications that can also make use of HMM outputs. As for our next step, we would explore the following directions: 1) through detecting patterns in search process, the system can predict which type of search tactic a user is employing and should employ in the next step. Through this way, the system can provide support intelligently; 2) it can be used to evaluate a search system works based on how well it supports the particular transition process and to identify how to make the transition smoothly and thus increase the performance.

DISCUSSION AND CONTRIBUTION

Through the analysis of search processes in exploratory search using HMM, we have several important findings. First, two types of hidden states are recognized in both individual and collaborative search processes: the search related hidden states and the sense-making related hidden states. This is compatible with previous studies that exploratory search can be viewed as an intertwine process of search and sense-making [28]. Second, using HMM in analyzing the individual search process and the collaborative search process, we obtained similar transition patterns as defined in several well-established information seeking process models, which demonstrates the validity of our model. Third, we found that the search-related states are similar in both individual and collaborative search, but sense-making related states are more complex in collaborative search. This again is consistent with previous findings. The ability to capture different types of sense-making demonstrates the generalizability of our model.

In terms of task type, collaborative search is more sensitive to the task difference. We didn't find many differences on the search process between two tasks in the individual search. However, the search processes between the two tasks in collaborative search are quite different. There are more transitions between search and explicit communications in the utility-based decision-making task. It indicates that this type of task requires users to be more active in the collaboration, which might be caused by the need for negotiation and agreement.

Based on the analysis of relationship between search process and search performance, we found that the transition from search to sense-making has a negative relationship with the performance while the transition from sense-making to search has a positive relationship with the performance. Here we refer to correlation not casual effects. We are not arguing that low performance is caused by the transition to sense-making. Instead, we think that when sense-making is needed, it means the user has to spend time on absorbing information or on resolving problems. If we can make the transition less frequent and smoother, it may improve the search performance. For example, a summary of workspace in the search frame may be able to help users to make sense of current search status without breaking the current search activity. In this case, users need less effort on sense-making so that they can keep searching without being interrupted. In future studies, we are planning to test whether the easiness of sense-making can help to improve collaborative search performance.

Our contributions include: 1) we demonstrated that HMM is a valid method for analyzing collaborative search process. The benefit of HMM is that it can model users' intent or hidden search tactics without input from a theoretical model or manual label. Collaborative Web search is an ideal case for applying HMM since collaborative search tactics have not been well-defined in literature. HMM can be useful for analyzing search processes that involve complex user interactions, such as e-discovery or multimedia retrieval; 2) we explored and identified the important procedures for applying HMM. We found that categorizing user interactions and conducting model selection are two important components in HMM. Our approach of categorizing user interactions using three dimensions: method, object and source can be applied to other studies. Future applications of HMM can follow the rules and methods we provided; 3) Comparing to the previous models, HMM provides more detailed information of transition probabilities among different states in the search process. Our study demonstrated that such information can be used for analyzing task differences and correlation with search performance. The findings can be used for evaluating collaborative Web search systems and guiding the design of the system.

We do acknowledge some limitations of this study. First, our search performance measures mainly consider topical relevance, whereas other measures might be needed for the utility-based decision-making task. This is because users' personal judgment matters for the relevance in this kind of tasks rather than the topical relevance. Second, we found that the recall is significantly lower in decision-making tasks. One reason is that the goal of this type of task is not to exhaust relevant information. Furthermore, we defined the relevant documents as those were saved by at least two teams, which might not be a good indicator of relevance in utility-based decision-making tasks. In the future, we will explore measures for more accurate reflection of search performance in collaborative Web search.

CONCLUSION

In this paper, we applied a novel approach for modeling search process automatically using hidden states. A user study was conducted to compare the search process in collaborative exploratory search and individual exploratory search. Through the analysis, we demonstrated that HMM is a valid method for automatically analyzing search processes. Different patterns of hidden states were recognized and compared in both individual and collaborative search. In addition, the patterns of hidden states between two types of tasks were quite different in collaborative search. We also discovered the relationships between search processes and search outcomes. Future works will focus on the application of these findings to facilitate collaborative exploratory Web search.

REFERENCES

[1] Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 19, 6 (1974), 716–723.

[2] Amershi, S. and Morris, M.R. 2008. CoSearch: a system for colocated collaborative web search. *Proceeding of ACM SIGCHI* (Florence, Italy, 2008), 1647–1656.

[3] Bates, M.J. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online Review*. 13, (1989), 407–424.

[4] Belkin, N. 1995. Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Information Retrieval Systems. *Expert Systems with Applications*. 9, 3 (1995), 379–395.

[5] Cao, H. et al. 2009. Towards context-aware search by learning a very large variable length hidden markov model from search logs. *Proceedings of the 18th international conference on World wide web WWW 09*. 11, (2009), 191.

[6] Chen, H.-M. and Cooper, M.D. 2002. Stochastic modeling of usage patterns in a web-based information system. *Journal of the American Society for Information Science and Technology*. 53, 7 (2002), 536–548.

[7] Chen, H.-M. and Cooper, M.D. 2001. Using clustering techniques to detect usage patterns in a Web-based information system. *Journal of the American Society for Information Science and Technology*. 52, 11 (2001), 888–904.

[8] Dervin, B. et al. 2003. *Sense-Making methodology reader: Selected Writings of Brenda Dervin*. Hampton Press, Inc.

[9] Diriye, A. et al. Querium : A Session-Based Collaborative Search System. 2–3.

[10] Ellis, D. 1993. Modeling the information seeking patterns of academic researchers: a ground theory approach. *Library Quarterly*. 63, 4 (1993), 469–486.

[11] Elsweiler, D. et al. Evaluating Personal Information Management Behaviour using Markov Models. 8, 280–288.

[12] Evans, B.M. and Chi, E.H. 2008. Towards a model of understanding social search. *Proceeding of the 2008 ACM workshop on Search in social media - SSM '08*. (2008), 83.

[13] Fuhr, N. 2008. A probability ranking principle for interactive information retrieval. *Information Retrieval*. 11, 3 (Feb. 2008), 251–265.

[14] Golovchinsky, G. et al. 2009. Collaborative information seeking. *IEEE Computer*. 42, 3 (2009).

[15] Golovchinsky, G. and Adcock, J. 2008. Cerchiamo: a collaborative exploratory search tool. *Proceedings of the computer supported cooperative work*. (2008).

[16] Hearst, M. 2009. *Search User Interfaces*. Cambridge University Press.

[17] Holscher, C. and Strube, G. 2000. Web search behavior of Internet experts and newbies. 33, (2000), 337–346.

[18] Hyldegard, J. 2006. Collaborative Information Behavior - Exploring Kuhlthau's Information Search Process Model in a Group-based Educational setting. *Information Processing & Management*. 42, 1 (2006), 276–298.

[19] Kim, J. 2009. Describing and Predicting Information-Seeking Behavior on the Web. *Journal of the American Society for Information Science and Technology*. 60, 4 (2009), 679–693.

[20] Kuhlthau, C.C. 1991. Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*. 42, 5 (Jun. 1991), 361–371.

[21] L.E.Baum et al. 1970. A maximization technique occuring in the statistical analysis of probabilistic

functions of Markov chains. *Ann. Math. Statist*. 41, 1 (1970).

[22] Marchionini, G. 1995. *Information seeking in electronic environments*. Cambridge University Press.

[23] McQuarrie, A. and Tsai, C. 1998. *Regression and Time Series Model Selection*.

[24] Morris, M.R. 2008. A survey of collaborative web search practices. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. (2008), 1657.

[25] Morris, M.R. 2013. Collaborative search revisited. *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13* (New York, New York, USA, Feb. 2013), 1181.

[26] Morris, M.R. and Horvitz, E. 2007. *SearchTogether: An interface for collaborative web search*. ACM Press.

[27] Paul, S.A. and Morris, M.R. Sensemaking in Collaborative Web Search. *Social Cognition*. 1–66.

[28] Qu, Y. and Furnas, G. 2007. Model-driven formative evaluation of exploratory search: A study under a sensemaking framework. *Information Processing & Management*. 44, (Nov. 2007), 534–555.

[29] Russell, D.M. et al. 1993. The cost structure of sensemaking. *Proceedings of the SIGCHI conference on Human factors in computing systems CHI 93* (1993), 269–276.

[30] Saracevic, T. 2007. Relevance : A Review of the Literature and a Framework for Thinking on the Notion in Information Science . Part II : Nature and Manifestations of Relevance *. 30, October (2007), 1915–1933.

[31] Shah, C. and González-Ibáñez, R. 2011. Evaluating the synergic effect of collaboration in information seeking. *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information - SIGIR '11*. (2011), 913.

[32] Shah, C. and González-ibáñez, R. 2010. Exploring Information Seeking Processes in Collaborative Search Tasks Chirag Shah. *ASIS&T* (2010).

[33] Shah, C. and Marchionini, G. 2010. Awareness in Collaborative Information Seeking. *Journal of the American Society for Information Science and Technology*. 61, 10 (2010), 1970–1986.

[34] Spink, a 2002. A user-centered approach to evaluating human interaction with Web search engines: an exploratory study. *Information Processing & Management*. 38, 3 (May. 2002), 401–426.

[35] Stebbins, R.A. 2009. Leisure and its relationship to library and information science: Bridging the gap. *Library Trends*. 57, (2009), 618–631.

[36] Tran, V.T. and Fuhr, N. 2012. Using eye-tracking with dynamic areas of interest for analyzing interactive information retrieval. *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval (SIGIR '12).* (2012).

[37] Wang, K. and Gloy, N. 2010. Inferring Search Behaviors Using Partially Observable Markov (POM) Model. *Web Search and Data Mining* (2010), 211–220.

[38] Xie, I. and Joo, S. 2010. Transitions in Search Tactics During the Web-Based Search Process. *Journal of the American Society for Information Science*. 61, 11 (2010), 2188–2205.