



Implicit search feature based approach to assist users in exploratory search tasks



Chathra Hendaheva^{a,*}, Chirag Shah^b

^a Department of Computer Science, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA

^b School of Communication and Information, Rutgers University, 4 Huntington Street, New Brunswick, NJ 08901-1071, USA

ARTICLE INFO

Article history:

Received 5 December 2014

Received in revised form 26 May 2015

Accepted 1 June 2015

Available online 18 June 2015

Keywords:

Exploratory search

Evaluation

User behavior

Implicit features

ABSTRACT

Analyzing and modeling users' online search behaviors when conducting exploratory search tasks could be instrumental in discovering search behavior patterns that can then be leveraged to assist users in reaching their search task goals. We propose a framework for evaluating exploratory search based on implicit features and user search action sequences extracted from the transactional log data to model different aspects of exploratory search namely uncertainty, creativity, exploration, and knowledge discovery. We show the effectiveness of the proposed framework by demonstrating how it can be used to understand and evaluate user search performance and thereby make meaningful recommendations to improve the overall search performance of users. We used data collected from a user study consisting of 18 users conducting an exploratory search task for two sessions with two different topics in the experimental analysis. With this analysis we show that we can effectively model their behavior using implicit features to predict the user's future performance level with above 70% accuracy in most cases. Further, using simulations we demonstrate that our search process based recommendations improve the search performance of low performing users over time and validate these findings using both qualitative and quantitative approaches.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Analyzing people's online search behaviors has become a major topic of interest within the domain of Information Retrieval as well as within the search engine industry in order to improve search and retrieval features. Examples of some recent studies include providing personalized search results based on geographical location and user profiles (Bennett, Radlinski, White, & Yilmaz, 2011; Sugiyama, Hatano, & Yoshikawa, 2004), re-ranking search results based on user's past search behavior (Matthijs & Radlinski, 2011) and providing trending news and stories based on different factors (Al Bawab, Mills, & Crespo, 2012) using services such as 'Google trends' and 'Yahoo Trending now'. Since the early 90s when the World Wide Web (WWW) was introduced, information seekers have greatly benefited from these technologies and the vast amount of information that could be accessed, which was a rarity up until that point in time. With the advent of Web search engines, users became accustomed to using a search engine to retrieve relevant information with keyword-based searches without having to sift through the vast amount of information repositories available on the World Wide Web.

* Corresponding author.

E-mail address: chathrah@cs.rutgers.edu (C. Hendaheva).

In any Web search there are two major constituents. One is the search system and the other is the user. User and system interaction has been explained in the stratified interaction model by [Saracevic \(1997\)](#). In an online Web search environment, usually the system is a Web search engine. Since each user is unique and the way people search varies depending on different factors such as information need, demographics, previous search experience, knowledge of the topic being searched on, cognitive abilities, personal traits, etc., understanding the user is a vital first step towards providing a personalized search experience. Understanding the user can be accomplished by building a user behavior model through observing his/her search activities and constructing a user profile to thereby make predictions and personalizations that help the user find relevant information.

In the domain of user behavior modeling, past research has focused on deriving user behavior models and then building sequential models to capture behavior patterns. [Fox, Karnawat, Mydland, Dumais, and White \(2005\)](#) described user behavior patterns as sequences and identified correlations between behavior patterns and user satisfaction. [White and Drucker \(2007\)](#) investigated how different user behavioral sequences can be linked to different types of Web users such as navigators and explorers. They identified that users who exhibited consistent behavioral sequences among different search sessions were mostly *navigators* while users with variable behavioral sequences were considered *explorers*. This type of user categorization based on behavioral patterns exhibited in the user search logs gave rise to interesting dimensions of modeling user profiles that could be leveraged to personalize and also assist in the user's search.

Another interesting analysis was performed by [Wildemuth \(2004\)](#) to examine the search tactics of users when searching for information. Common search tactics were found among the study participants. It was also observed that tactics changed over time during a longitudinal study depending on the level of domain knowledge accrued using state transition matrices and maximal repeating patterns techniques.

There has been some notable work on temporal modeling of user search behavior in order to capture how users interact with information and the system while conducting information seeking tasks. [Azzopardi \(2014\)](#) analyzed user search behavior by introducing a new economic theory-based model where the different interactions that could be performed by the users on the search interface were analyzed as a cost/gain model. This model considered eight interaction hypotheses pertaining to search interactions with the system and were found to be in line with past empirical observations and the expected user behaviors. [Smucker and Clarke \(2012\)](#) evaluated the search process more effectively using a time-based gain measure. Their work focused on assessing how a user evaluates a search results page over time and modeling the gain by viewing a document ranked at rank k . Their work provides a better evaluation framework for document ranking by accommodating aspects of the search process such as time, decay, and gain that are ignored by traditional effectiveness measures. Further work on search results relevance personalization was performed by [Bennett et al. \(2012\)](#) by analyzing searchers' short-term (session) and long-term (historic) behaviors. They found that substantial benefits could be gained in search relevance at the beginning of a session using historic behaviors while session behavior contributes mostly in an extended search session. Another interesting finding is that the combination of both session and historic behavior outperforms using either factor in isolation in providing relevant documents, thus enhancing search personalization.

Although, there has been work done to build meaningful user behavioral models, to the best of our knowledge there is a lack of research regarding the development of user behavioral models specifically for exploratory search tasks. Given that exploratory search is open ended, multi-faceted, and lacks clear boundaries, the way users behave during an exploratory search task is highly variable. The main aspects of exploratory search that make it complex and unique compared to other types of information searches such as known-item and fact-finding searches are uncertainty, creativity, innovation, knowledge discovery, investigation, and learning ([Bawden, 1986](#); [Marchionini, 2006](#); [Foster & Ford, 2003](#)). Also, since such tasks tend to extend over longer time spans or even multiple sessions, the importance of user search assistance and direction to improve the user experience and user goal achievement becomes paramount.

Therefore, providing an approach one could use to reasonably analyze the behavior patterns of a user when performing an online exploratory search task would be extremely helpful in identifying how a user is performing. This would also provide much better recommendations and a more personalized information seeking experience, which in turn results in better user satisfaction and achievement of search goals.

In this article, we develop a framework to understand the underlying search process by using the user's search actions and implicit features to analyze exploratory search through a quantitative approach. Explicitly we try to address the following research questions.

- **RQ a:** How effectively can we model different aspects of exploratory search such as knowledge discovery, creativity, exploration using implicit features extracted from the user search process to identify underperforming users?
- **RQ b:** How effectively can we provide search process based recommendations to improve user search performance?

2. Background on exploratory search

Analyzing a user's search behavior when performing online search tasks could be used to discover insights that might enable one to assist the searcher in a more successful search experience. Out of the many types of online information seeking tasks that occur every second on the Web, this article only focuses on exploratory search tasks that usually require longer time spans with multiple search queries. Exploratory search is a specialization of information seeking, which describes the activity of attempting to obtain information through a combination of querying and collection browsing ([White & Roth,](#)

2009). Further, it is a search task with multiple facets, multiple queries, or even multiple search sessions. This allows the information searcher to find meaningful and useful information to extend a user's knowledge about the task at hand.

What makes exploratory search *exploratory* is that this type of information seeking requires users to undergo complex cognitive tasks that lead to learning, exploring, and acquiring, intellectual skills. The searching and browsing behavior of information seekers enables the searchers to learn and develop skills while conducting exploratory search tasks. Users conducting exploratory searches are generally (1) unfamiliar with the domain of their goal (i.e., need to learn about the topic in order to understand how to achieve their goal); (2) unsure about the ways to achieve their goals (either the technology or the process); and/or even (3) unsure about their goals (White & Roth, 2009).

Most of the research conducted in the area of exploratory search has focused on building special exploratory search systems that help searchers conduct such tasks in an effective manner (White, Muresan, & Marchionini, 2006). Some examples of prominent exploratory search browsers are Flamenco (Yee, Swearingen, Li, & Hearst, 2003), mSpace (schraefel, Wilson, Russell, & Smith, 2006) and Relational Browser++ (Zhang & Marchionini, 2005) that support various levels of faceted browsing in addition to traditional keyword search capabilities. These browsers provide an enhanced experience for users looking to navigate a larger corpus of data in an organized manner.

With the advent of such systems and enhanced browser capabilities that support exploratory search, the major emphasis has been placed on system-based evaluation to identify which systems are suitable for different search conditions. In such system-focused evaluations, factors such as the system utility in measuring how successful the system is in facilitating the key elements of search exploration, recall, and providing relevant documents based on user learning, are considered. Wilson, schraefel, and White (2009) provides a detailed inspection framework for the evaluation of such search interfaces to identify their strengths and weaknesses and to evaluate their applicability to support user search tactics and conditions when conducting a search.

Although building new systems/tools or enhancing the existing system/tools for exploratory search by evaluating the system performance is of importance, it is also crucial to evaluate how an online information searcher is performing irrespective of a specific exploratory search system/tool. Focusing on the user's search process would lead to a better understanding of searcher strategies and underlying intentions that are not captured in a system setting.

Borgman, Hirsh, and Hiller (1996) stated that "The ability to study online searching behavior in all its complexity requires research methods that can capture the details of the search process, rather than just the product or output of a search session". They suggested that building quantitative models of a user's search process based on transactional log data would be beneficial in capturing user search behavior. Information theoretical models such as the ISP model by Kuhlthau (2004) describe the process of information seeking to accomplish a task goal from a user's perspective with regard to six stages: initiation, selection, exploration, formulation, collection, and presentation. The types of search behaviors that are exhibited within each of these stages of the ISP model tend to vary in describing how the search process changes with time while conducting the search task. Vakkari (2001) developed a study to expand the ideas based on the ISP model by conducting a longitudinal study with student participants preparing a research proposal related to information search. These research findings revealed that it is productive to study information searching and information retrieval as a process in connection with the search task that generates that process. Further work by Liu and Belkin (2012) to analyze how search systems could help in search tasks found that although the task type and task session did not affect users task performance, their familiarity with the topic and experience level of conducting that task had an influence on users' search performance.

It is important to formalize an evaluation framework for exploratory search that is user-centric, focusing on the main aspects of exploratory search such as user learning, knowledge discovery, creativity, engagement, and investigation. One attempt to shift the focus of evaluating exploratory search to a user-centric analysis was done by Spink (2002) with the use of a new search tool covering aspects of effectiveness and usability by taking a qualitative approach. Another approach that is highly relevant to capturing and studying exploratory search behavior is information foraging theory (Pirulli & Card, 1999). With this theory, Pirulli and Card showed that using information cues that provide orientation, navigation, assessment, and information value are helpful to the user for making informed decisions. The authors showed users also estimate how much information they are likely to retrieve from a given search path and after comparing actual outcomes with their predictions (information scent) and when no added value is incurred for the cost of searching, they move on to other information sources. Nevertheless, the information foraging approach suggests how to optimize the information shown to the user such as Web pages and relevant snippets along with Web links to help users in their information seeking process. It has also been identified by Singh, Hsu, and Moon (2013) that it is essential to capture and represent the variability in the semantics of the information that may be relevant to a query and also to support efficient and effective interaction between users and the information even when the search goal is not well-defined early in the search process. The authors explained that such an approach requires user-centric capabilities rather than pure-syntactic query-retrieval to facilitate exploratory search. Their approach constructs a presentation-interaction paradigm based on content-pages correlations between Web pages to extract different information and perceptual cues to show results to users with varied perspectives in order to facilitate exploratory search. Further Ruthven (2003) found that only certain query expansions can lead to effectiveness where searchers use them to expand their search. Thus, past research has mainly focused on optimizing the results for query-based retrieval and showing those results in an improved manner to the users to facilitate exploratory search rather than analyzing the actual user-centric search process of when and how users got to that information.

There is a gap in this research area of dynamic user-centric search evaluation for exploratory search that requires further investigation. Given that most current online searchers use search engines like Google, Bing, and Yahoo! for their day-to-day

search tasks, it is extremely useful to understand their search performance when conducting an exploratory search task using easily available search engines. Evaluating the user search performance dynamically by analyzing search behavior over a period of time while making the user aware of their performance level would be an immense contribution to users as well as to the community of search engine developers and other exploratory search systems/tools. Therefore, a user-centric exploratory search evaluation framework would enable users to be more proactive and effective when conducting search tasks. This would ultimately lead to a better user experience and higher user satisfaction.

3. Proposed approach

The proposed approach attempts to build a framework to analyze the search behavior of individuals who are performing exploratory search tasks using a user-centric search process based approach. We try to incorporate different aspects that define exploratory search when evaluating the user search process over time and also apply user search behavior patterns in understanding the underlying process that guides the exploratory search. Since the underlying search process is unique for each individual, this proposed approach would try to provide a more personalized understanding and evaluation of the user search performance.

In order to facilitate this approach, we propose to analyze the search process, which can be considered as a collection of actions a typical user would take when performing an online search task, including but not limited to the following.

- Issue a query.
- Scan the search result pages.
- Visit a particular Web page.
- Bookmark or save the page for future use.
- Reformulate a query.
- Revisit a Web page.
- Collect a passage from a Web page.

This approach would consider the search process as a sequence of actions a user performs when conducting an online search task. This would provide a means to understand *why* and *how* a user conducts a search as well as indicate whether that would lead to success in achieving his/her goal.

We base our framework on the assumption that the actions and behaviors that users exhibit via different actions and decisions they make are results of the underlying cognitive process they undergo when conducting the search task.

4. Methodology

The two major parts of the proposed approach are capturing user search behavior as sequences of actions and analyzing the time series of features that help evaluate aspects of exploratory search. This section explains the above two components of our proposed approach in detail. We then provide an explanation of how we are using these features and sequences of actions to understand and evaluate user search performance when conducting exploratory search tasks.

4.1. Sequence analysis and segmentation

In order to analyze the search processes followed by different users, we assume that the underlying dynamics of the such processes are expressed by a collection of actions that take place from the beginning to the end of the search episodes.

In order to segment a user's online search process, we represent/annotate each user search process as a sequence of actions. For this exploratory search task study, the users performed the following types of actions.

- Issuing a query.
- Visiting a Search Engine Result Page (SERP).
- Visiting a Web content page.
- Collecting snippet from a Web page.
- Viewing collected snippets.
- Viewing previously visited site.
- Re-issuing a query.

Following are the steps involved in analyzing the online search process based sequences. We follow the same approach proposed by [Hendahewa and Shah \(2013\)](#) and [Cole, Hendahewa, Belkin, and Shah \(2014\)](#) in analyzing search activity sequences.

- *Step 1 – Sequence construction:* The above action sequences can be constructed for each individual user for the observable time from task inception to task completion. Each action was considered at the atomic level of seconds, so if a user

continues looking at a Web content page for t mins, then the number representing the Web page visit action was repeated $t * 60$ in constructing the user's overall sequence.

- **Step 2 – Sub-sequence construction:** Once we have the user search process sequences, it is essential to use sub-sequences derived from user search process sequences in order to analyze how similar or different each sub-sequence is from the rest. If we analyze the whole sequence at once, the similarity measurement will not be valid since the patterns in sequences change over time. In order to overcome this, we can adapt the method of sliding windows, which is widely used in time series analysis. We can use sliding window with different lengths of w minutes of the overall sequences with a gap of m minutes to come up with different sets of sub-sequences.
- **Step 3 – Similarity calculation:** Finding the similarity between the sub-sequences is essential to identifying which sub-sequences have high similarity compared to others within and among users. Hamming distance is one such distance calculation method applicable here.
- **Step 4 – Clustering:** Next, we use a clustering method such as hierarchical agglomerative clustering to cluster the sub-sequences and find an effective number of clusters, which would represent the user search process based stages.

It is also important to analyze how these sub-sequences, which were grouped into different clusters, were positioned along the search process. After sub-sequence clustering, we map each of the sub-sequences to their corresponding cluster within its original position in the user search process.

4.2. Exploratory search behavior analysis

As explained in Section 2, exploratory search relates to complex searches that involve an element of uncertainty. This distinguishes exploratory searches from other types of information seeking such as fact-finding or known-item searches. Some of the major constituents of exploratory search are uncertainty, creativity, innovation, knowledge discovery, learning, and investigation. In this approach, we attempt to capture most of the above mentioned aspects of exploratory search using implicit measures that we observe when users conduct exploratory searches. These elements can be captured without explicitly asking users questions about their search process.

In order to understand behaviors exhibited in the search process over time we extract different implicit features from a user search process while he/she is conducting an exploratory search task. Most of these features have been used in other settings in modeling user behavior in collaborative search in previous work (e.g., Shah & González-Ibáñez, 2011, 2012) but, not in the setting of an exploratory search evaluation framework. Shah, Hendahewa, and González-Ibáñez (2015) has successfully shown that one could predict the user search performance along the search process for both individuals and dyads using user search process based implicit features extracted from the transaction log data. We are incorporating those implicit features in addition to a few more in this article to capture the underlying search process of individual users conducting exploratory search tasks. During online information searching, users interact mainly with Web pages and query using different search actions. Therefore, we are using implicit features that capture the search process along those two artifacts that span across multiple aspects of exploratory search mentioned below. These provide insights into how each individual user performs throughout the search process to achieve their task goal.

- **Uncertainty** – Given the definition of exploratory search from a user perspective, such that it originates from having no clear idea about the goal or how to achieve it, it shows that the exploratory search is heavily vested in uncertainty. Affective and cognitive uncertainties are persistent characteristics of exploratory search based information seeking (White & Roth, 2009, pp. 10). Uncertainty is a natural user experience within the process of information seeking and acquiring meaning that can give rise to feelings of doubt, confusion, frustration, and anxiety (Kuhlthau, 2004) that can be identified based on the type of actions that users perform over the search process. Therefore, we model uncertainty based on the patterns of various action sequences over time.
- **Exploration** – As the name implies *exploration* is the most important constituent in exploratory search. As White and Roth defined it (White & Roth, 2009, pp. 6) “Information exploration is a broad class of activities where new information is sought in a defined conceptual area”. The idea of information being sought and new information being found can be modeled by observing the search process of a user over time. Features pertaining to information coverage based on Web pages and information seeking based on query execution can be used to identify when and how the user found information leading to exploration.
- **Creativity** – This aspect plays a significant role in exploratory search since there is no clearly defined search path to reach the goal. There can be multiple search paths and multiple search activities to achieve the desired outcome even for the same search need or goal. Therefore, the user's creativity in how he/she conducts the search process has a major impact on the entire search process. This could be identified by observing how creative and innovative a user is when seeking information by capturing how much novel information they are able to find and the level of difficulty required to locate that information.
- **Knowledge discovery** – When a user conducts an exploratory search, he/she gathers knowledge about the search topic/area. It enables users to fill an existing gap in knowledge on a specific topic/area. As a result of exploration, the search process exposes users to content that helps relate the problem context to similar documented experiences and promote

information discovery. These can be operationalized by observing implicit features pertaining to relevant and useful information discovery that leads to knowledge acquisition.

Table 1 summarizes the different aspects of exploratory search that are being considered in order to quantify and evaluate the search process of users, as mentioned above.

4.3. Feature extraction

As indicated in Table 1, the following features were extracted corresponding to capturing different aspects of exploratory search.

- Complexity of action sequences (u, t) – $C_{u,t}$: In order to find the complexity of the sequences (s) for each user u up to time t , we used the complexity index (C) as described below. The complexity index (Gabadinho, Ritschard, Studer, & Muller, 2010) is a composite measure that combines the number of transitions in the sequence with the longitudinal entropy. This shows us the different levels of action sequences that users enact in conducting the task. When we analyze this over time it provides a snapshot of the sequences' complexity and how the sequences change throughout the search process, providing some insight into user behaviors.

$$C_{u,t} = \sqrt{\frac{q(s) h(s)}{q_{max} h_{max}}} \quad (1)$$

where $q(s)$ is the number of transitions in sequences s , q_{max} is the maximum number of possible transitions, $h(s)$ is the within entropy and h_{max} is the maximum entropy calculated as $\log a$ where a denotes the alphabet of the sequence.

- Universe of coverage – U : The universe of distinct content pages that all users u for the same topic when doing the exploratory search task.
- Unique Web pages (u, t) – $UW_{u,t}$: Number of content Web pages only visited by user u up to time t out of the Universe of Coverage for that topic U . This feature shows the novelty of information found by each user.
- Likelihood of Discovery (u, t) – $LD_{u,t}$: It is important to understand how difficult it was to discover the information that each user located. Shah and González-Ibáñez (2011) referred to this as Likelihood of Discovery (LD), and it is defined as follows.

$$LD_{wp_i} = \frac{-1 \cdot n\{wp_i\}}{|U|} \quad (2)$$

Therefore the discovery measure for each user over time can be found as follows.

$$LD_{u,t} = \frac{\sum_{i=1}^{|Coverage_t|} LD_{wp_i}}{|Coverage_t|} \quad (3)$$

- Coverage – $Coverage_{u,t}$: The set of distinct content pages that user u visited up to time t .
- Total Coverage (u, t) – $TC_{u,t}$: The total number of distinct Web pages visited by a user (u) up to time t . This feature captures the Web page based activity performed by a user and provides a measure to see how much distinct information has been found by the user up to a particular time.
- Number of Distinct Queries – $NQ_{u,t}$: The number of distinct queries issued by the user (u) up to time t .
- Useful Coverage (u, t) – $UC_{u,t}$: The total number of distinct Web pages on which a user spent at least 30 s, up to time t . This measure evaluates out of the total pages the user visited how many of them were useful in finding relevant information leading to satisfaction with their context in completing the exploratory search task.

Table 1
Exploratory search analysis criteria.

Major aspects of exploratory search	Evaluation type	Implicit features
Uncertainty	Level of frustration Anxiety Fluctuations in activity behavior	Complexity of action sequences
Creativity	Information novelty Difficulty to find information	Unique Web pages Likelihood of discovery
Exploration	Exploration of information Web pages Queries	# of content pages visited # of different queries issued
Knowledge discovery	Finding useful information Finding relevant information	Useful coverage Relevant coverage

- Relevant Coverage (u,t) – $RC_{u,t}$: The total number of distinct Web pages from where user has collected snippets (information copied from a Web page) from up to time t . This measure evaluates how much information was considered relevant to that specific user from what he/she found in completing the task.
- Number of Saved Snippets (u,t) – $NS_{u,t}$: Total number of snippets saved by user u up to time t . This measures the amount of information that the user thought might be relevant in the future to complete the task and needed to be remembered. In other words, this feature is an indication of explicit relevance judgments made by the user.
- Coverage Effectiveness (u,t) – $CE_{u,t}$: Coverage Effectiveness of user u up to time t is defined as the ratio of useful coverage and total coverage.

$$CE_{u,t} = \frac{UC_{u,t}}{TC_{u,t}} \quad (4)$$

- Query Effectiveness (u,t) – $QE_{u,t}$: This implicit feature measures for each user u up to time t based on how much they managed to achieve by issuing queries. This measure was used in evaluating time-based user search performance.

$$QE_{u,t} = \frac{CE_{u,t}}{NQ_{u,t}} \quad (5)$$

- Recall (u,t) – $R_{u,t}$: We define recall for each user, u up to time t based on the universe of relevant pages collected by all users conducting the similar topic based exploratory task, U_r .

$$R_{u,t} = \frac{|RC_{u,t}|}{|U_r|} \quad (6)$$

- Precision (u,t) – $P_{u,t}$: We define precision for each user, u up to time t as follows.

$$P_{u,t} = \frac{|RC_{u,t}|}{|Coverage_t|} \quad (7)$$

Precision and recall based measures were used as the baseline features in the analysis.

4.4. Analysis method

User search sequence based features, such as sequence complexity, and time series based features can be combined to evaluate how each individual user is performing along the search process. This will form our experimental system. In order to build a baseline method by which to compare our proposed approach, we will measure against precision and recall-based time series in order to determine whether we can outperform these metrics by better predicting and understanding search behavior when the user is conducting the exploratory search task. Since we do not have a ground truth label of task success (since there is no single correct answer for exploratory search tasks), we would use an unsupervised learning method to group users into classes based on their observed behavior patterns and time series features. This grouping would then help us understand their different levels of performance over time. We then evaluate different levels of performance through aspects of exploratory search such as uncertainty, knowledge discovery, creativity, etc (described in Table 1) to make judgments regarding the user's performance level and compare them against other users who worked on similar topics in exploratory search tasks. We also identify the level of search behavior based on the sequence patterns and analyze the user's effectiveness in exhibiting the different behavior patterns such as exploration, investigation, and uncertainty. This method of precision and recall based evaluation would also provide the performance level scores over time that would occur if we did not observe user search sequence patterns nor analyze the time based implicit features.

5. Data

In this analysis we used transactional log data collected from a user study conducted when users performed an exploratory search task. The log data is comprised of type of action type, pages visited, queries executed, and timestamps of all events.

- User ID.
- Time-stamp.
- Action (query, visited URL, saved snippets).

Target participants in this study were undergraduate students who are highly fluent in the English language with intermediate to advanced online search and typing skills. We targeted undergraduates from majors such as Computer Science, Information Technology, Informatics, Nursing and Biological studies through open calls that were publicized through various email-lists. Through an online registration form, the participants were able to sign up for the study.

The user study consisted of conducting an exploratory search task (that involved finding information relevant to the topic and compiling an article) with two sessions covering two different topics of interest selected by each user. The user study system was designed and developed as a web application and a browser add-on that can be accessed online. The system facilitated automatic progression of log in, questionnaires, video tutorials on how to use the system, a practice session, pre-task quiz, topic selection, searching information online to complete the exploratory search task, post-task quiz, and questionnaires. Since the browser plug-in is customized for the Firefox browser, we instructed the participants to install Firefox before accessing the user study system. Each individual participant was paid \$20 after completing both the sessions.

5.1. Task design

In order to make the task as close as possible to an exploratory task, the following features were taken into consideration in the design of the user study. These characteristics represent important facets of exploratory search, which have been absent from previously conducted user studies (e.g., [Shah & González-Ibáñez, 2011, 2012](#)). It should be noted that the tasks across different sessions and topics are similar since the purpose is finding relevant information online and compiling a report. For example, [Qvarfordt, Golovchinsky, Dunnigan, and Agapie \(2013\)](#) used a search task of finding as many relevant documents as possible for a user study involving an experimental user interface (Querium) across six different search topics. The only variation here was providing users the option of selecting among a set of topics. This simulates a more realistic environment so that users can work on topics that are of interest to them, leading to better motivation for finding information. Therefore we can argue that given that the nature of the task is similar (irrespective of the topic), the search process adopted by users to find information using the online search and browser plug-in tool would be substantially similar in all cases.

- Exploratory search tasks with longer time limits (1–2 h) to facilitate longer time spans, multiple queries, user engagement, learning and exploration if users are focused.
- Flexibility in approaching the task
 - Flexible location: This user study was facilitated using a in-house developed Web application and a browser plug in (to capture Web pages, snippets, provide instructions, timers and access to an editor to type in the report) that can be installed in the Firefox browser in any computer running either a Windows or Mac operating system. The users had the flexibility to work on their own time from anywhere if they had access to the above computer requirements and Web access.
 - Flexible use of tools: There was no restriction imposed on search engines they could use while participating in this study. This facilitated freedom for the users to work on any search engine(s) they were comfortable with. This provided a more realistic environment rather than asking the users explicitly to use a specific search engine.
 - Allow flexible times: Participants could use the system any day or any time of day. Each session lasted minimum of 1 h and a maximum of 2 h.
 - Allow users to select the topic they want to work on: This enabled the freedom to work on a topic the users were interested in in order to facilitate user interests, needs, and preferences driven by the exploratory search.
- Capture level or prior knowledge on the topic and evaluate the knowledge gained at the end of the task: This is achieved by providing pre and post session quiz questions for each topic. Users had to answer the same set of multiple choice questions before and after the task (5 quiz questions) in order to capture their prior knowledge and post-task knowledge gained by performing the exploratory search task.

We picked five different topics and constructed descriptions to create simulated work-tasks, similar to what is suggested by Borlund and Ingwersen ([Borlund & Ingwersen, 2006](#)). We also used the guidelines provided by [Wildemuth and Freund \(2012\)](#) in designing exploratory search tasks to be close to everyday life information problems that are open-ended while targeting multiple facets. Topics and task descriptions in each area are shown below.

- **Technology:** “You are a journalist for a national newspaper. With the recent focus on security breaches and consumer data breaches, vulnerabilities in data and software have become an important topic of interest to a multitude of people and organizations ranging from government, companies, technical industry, to the general public. You have been asked to write an article about data and software vulnerabilities. Your article should include a basic introduction to the topic, along with past and present events. Your article should appeal to as many people as possible, including an unfamiliar, lay audience and people in affected businesses. It should cover different aspects of software vulnerabilities that are relevant to their daily lives. The article should also focus on the measures taken to minimize these risks at industry, governments and consumer levels.

When conducting this task, you should collect as much information as you can by searching online. You can use snippets to collect information that you deem useful for writing the report and copy them into your article. Nevertheless, the article should be well written with a proper flow so that the readers would be able to understand the context and get lots of information about this topic. Your article should be around 1000 words.”

- **Health and wellness:** “One of your close family friends talk to you about his current health and wellness issues to you and since he is not tech-savvy you are trying to assist him/her find useful information so that he can read the report you write within this task to get more information about his health and wellness requirements. Your family friend is a 30-year-old, type 2 diabetic male who wants to lose weight. He has Internet access and an iPhone but is not that tech-savvy. However, he is usually very busy with work and family and can only spare 3 h each week to exercise. He has asked for your help. Assemble a diet and exercise program for him, including its benefits and risks. This report should also cover aspects of possible applications he could use easily to monitor and control his wellness. When conducting this task, you should collect as much information as you can by searching online. You can use snippets to collect information that you deem useful for writing the report and copy-paste them into your article. Nevertheless, the article should be written with proper context to cater to your family-friend's needs and need to elaborate on the various options he has to be healthy. Your report should be around 1000 words.”
- **Environment and energy:** “You are planning to buy a new house within the next year and you deeply care about being a ‘green’ citizen using energy efficient methods be part of your life style. You have heard that using solar energy panels for heating can save a lot of money. Since you are a first time home buyer and currently have limited knowledge about solar energy for home heating, you would like to research online to find more about this topic and construct a report that you can refer to and show your family members when making the decision later on. You are generally interested in finding basic facts about solar energy installations, issues to be considered when deciding between conventional methods of home heating and solar panels. You also need to find rough estimates of the cost of installation, maintenance and also any other financial impacts. You should also consider looking at other measure to make an energy-efficient home and their costs and benefits. When conducting this task, you should collect as much information as you can by searching online. You can use snippets to collect information that you deem useful for writing the report and copy-paste them into your article. Nevertheless, the report should be written in a way that would be easier for you and your family members to later look at and get a clear and comprehensive idea about the solar energy installations at a new home and how it will impact their financial, life style and consumption patterns, along with any other factor you consider as useful in this context. Your report should be around 1000 words.”
- **Entertainment:** “You are a student who has volunteered to write an article about latest technological developments in movie making and production to the college art magazine. Movie industry has been adapting new technology such as 3D, avatars, graphics, animation, etc and this has led to development of specific types of movie categories such as animated movies. Find information about different technologies adapted by animated movie creators, producers and directors in designing such movies, the major motion picture companies involved in this process and major research labs working on developing software and hardware relevant to this area. Present the history of those technologies from early animations to most recent animation technologies. Also include details about how those techniques have been utilized in at least 5 animated movies developed in the recent past. When conducting this task, you should collect as much information as you can by searching online. You can use snippets to collect information that you deem useful for writing the report and copy-paste them into your article. Nevertheless, the article should be organized with sub areas that you find interesting and relevant to this topic so that the reader would find is easy to follow and learn about this topic. Your article should be around 1000 words.”
- **Art and history:** “You are taking a class in Art history this summer and as a class assignment you have to write a report about Leonardo Da Vinci's work. This report should focus on general facts about the artist and also some popular works by him. In addition to those, you are also supposed to find details about other artists who were working for the Medici family in Florence during the time Da Vinci was painting the Last Supper in Milan? Can you find two artists who probably knew Leonardo? Show evidence that supports your claim that they knew each other. Further, you should find information about different methods of painting used by the artist, specially painting methods such as ‘like smoke’ and explain how this style has led to development on a new painting style later on. When conducting this task, you should collect as much information as you can by searching online. You can use snippets to collect information that you deem useful for writing the report and copy-paste them into your article. Nevertheless, the report should be organized with sub areas that you find interesting should cover major works of Leonardo Da Vinci. Your report should be around 1000 words.”

6. Results

From the user study we were able to recruit 18 undergraduates to complete 2 sessions of the task with the selection of 2 different topics to work on in each session. This resulted in a total data set with 36 exploratory search task sessions.

The distribution of topics selected by users are shown in Fig. 1. Since only 2 users selected the topic on Environment and Energy, we had to eliminate those two users from the analysis due to sparsity of data. Therefore, it should be noted that the overall analysis is performed on the remaining 34 sessions of exploratory search.

With regard to analyzing the pre and post level knowledge of the topic, we found that except for 3 user sessions, in all other cases the post quiz performance (number of questions with correct answers) increased or remained at the same level as the pre quiz performance. This amounts to quiz performance remaining at same level for 41% of the user sessions while having increased post-level performance for 53% of the user sessions.

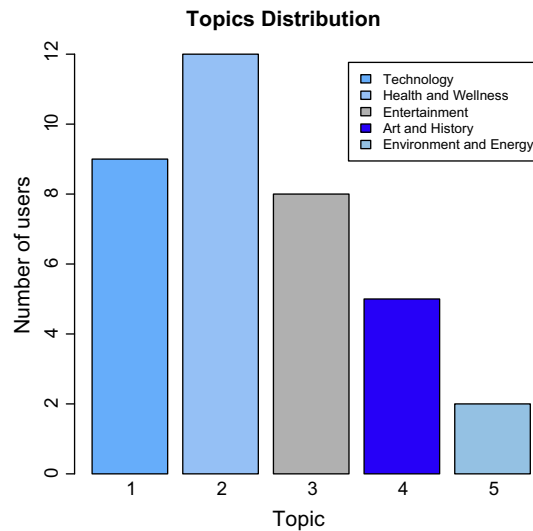


Fig. 1. Distribution of topics selected by users.

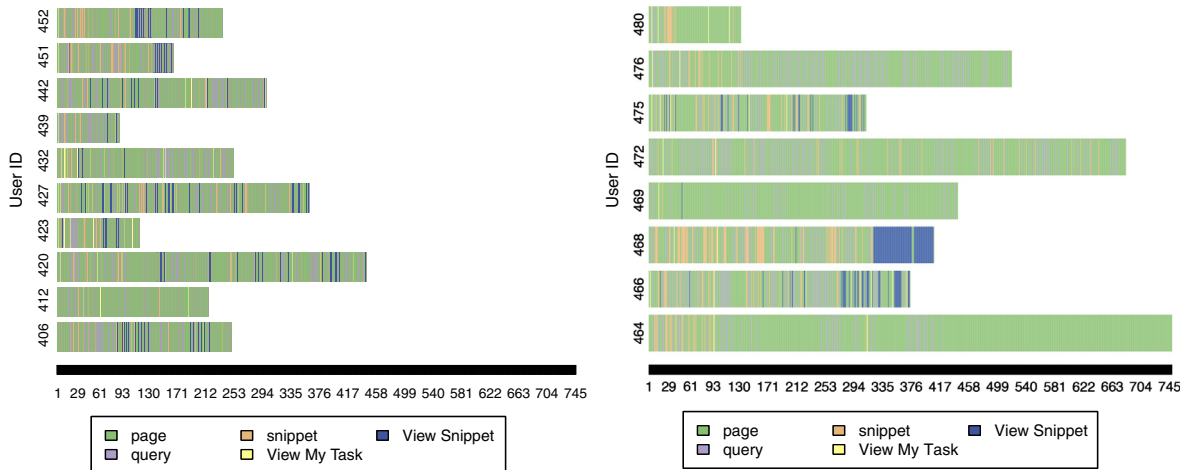


Fig. 2. Action sequences for users for Session 1.

The individual action sequences consisting of different user search actions (e.g., Issuing a query, visiting a Web page, collecting/viewing snippet) over time are depicted in Figs. 2 and 3. It can be observed that each user has highly variable actions sequences. Even the time participants spent on the exploratory task varies given that the task had flexible time limits as explained in Section 5. This illustrates the importance of analyzing the action sequences of users to discover meaningful insights from their behavior patterns.

For the user action sequences constructed from the search logs, the complexity measure was calculated to identify how different users and their actions can be considered as simple or complex as shown in Fig. 4. It can be observed within the same user the sequence complexity based on longitudinal analysis, session 1 and session 2 only varied slightly, except for few users. The difference between the complexity measures varied between the same user across session 1 and session 2 in the range of 0.0 to 0.2 for 88% of the users while only two users had change in complexity ranging from 0.25 to 0.3. This shows that most users exhibit a similar search behavior pattern when conducting the same type of exploratory search task even when the topics are varied.

In Fig. 5, we illustrate the time series of baseline features; precision and recall (calculated as in Eqs. (6) and (7)) for users who selected the topic on technology. It can be observed that the variation in precision over time is almost linear while recall has a slight increase over time.

Figs. 6 and 7 show the implicit features such as search action sub-sequence complexity (Eq. (1)), Query Effectiveness (Eq. (5)) and Coverage Effectiveness (Eq. (4)) of users within session 1 that illustrate how these features fluctuate over time.

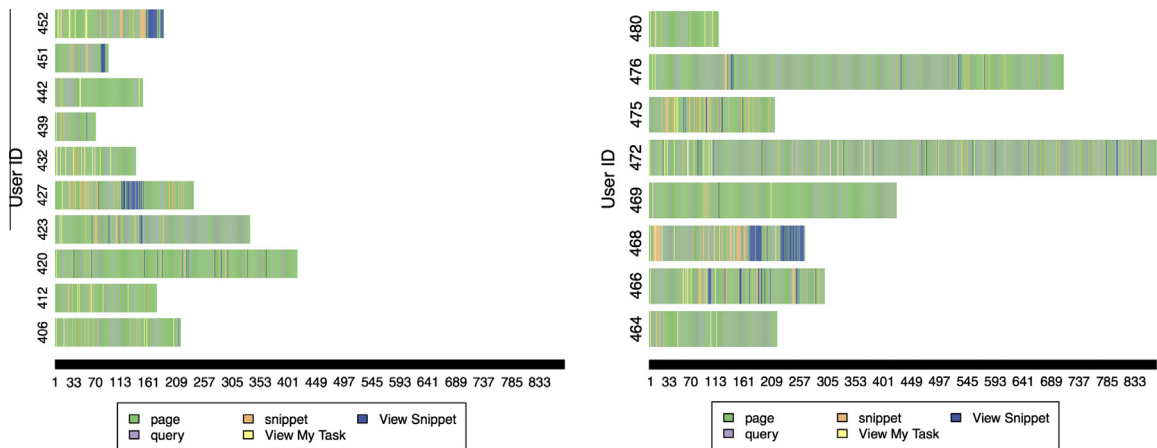
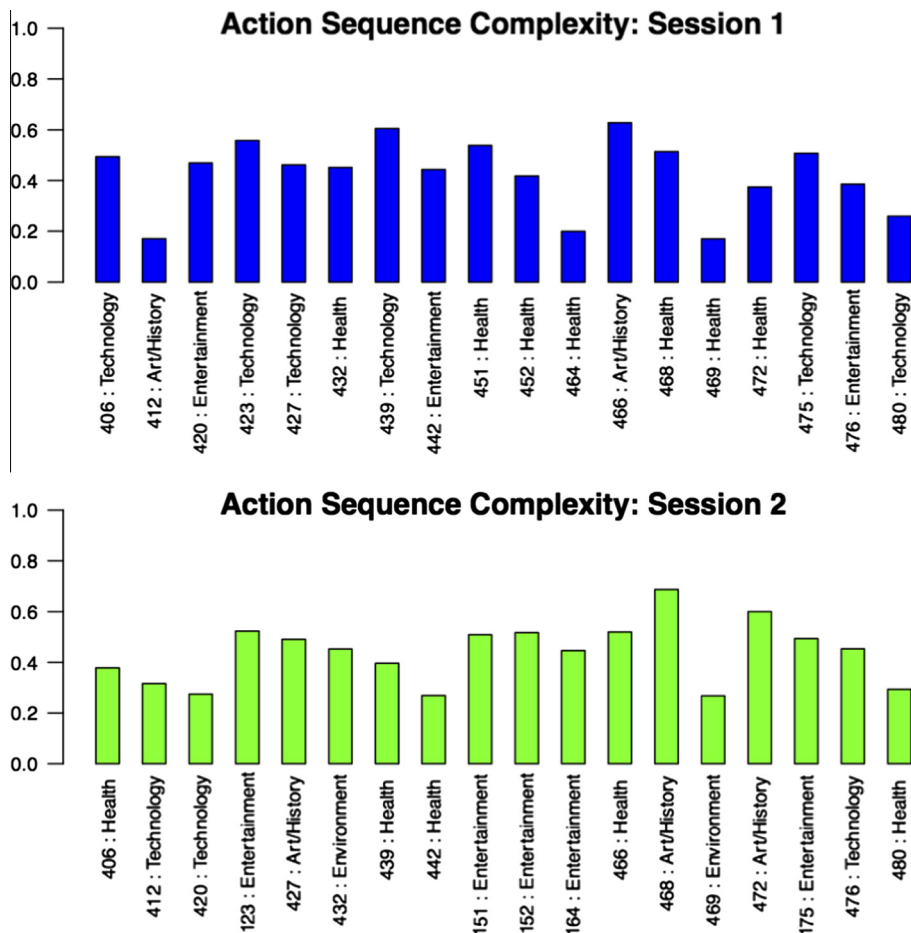


Fig. 3. Action sequences for users for Session 2.

Fig. 4. Complexity Index C_s for different user action sequences.

The baseline features such as precision and recall fail to capture the varying performance of the user search process unlike other features. These observations are in line with the findings and observations from previous research work such as those explained in Vakkari and Huuskonen (2012), Borgman et al. (1996), where the search process based indicators are better representations of search behavior than outcome based measures. Precision and recall show that they gradually increase over time, although other features such as query effectiveness, coverage effectiveness and sequence complexity, capture the

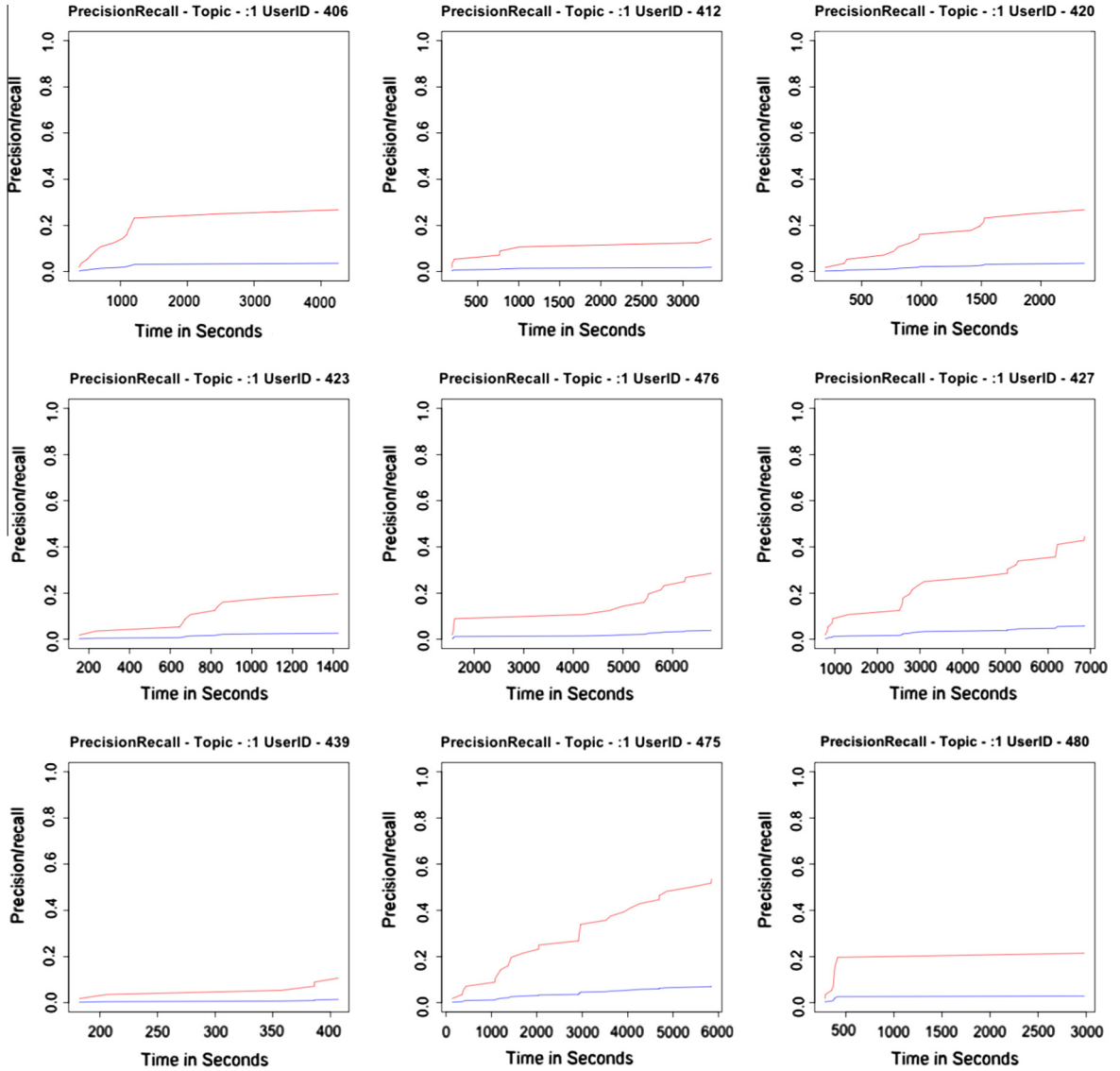


Fig. 5. Time-based baseline features for users for users who selected Topic 1: Technology.

underlying search process better by showing the variations in the search process over time. Baseline (Precision and Recall) based features are monotonically increasing over time but, search process and sequence based features are varied based on the actions the users perform. This captures the underlying search process more accurately.

6.1. User search performance evaluation

Using a binary classification model, the search performance for each user, u at each time t was classified based on the following criteria.

$$Class(u, t) = \begin{cases} 1 & \text{if } QE_{u,t} \geq \overline{QE_{u,t}} \\ 0 & \text{else} \end{cases}$$

where $\overline{QE_{u,t}}$ is the mean Query Effectiveness value for all users at time t .

We created feature matrices using the features extracted for each user for each time based on Table 1. We used 75/25 split of the data points for training and testing respectively. The training of the classification model was performed using Support Vector Machines (SVM) with kernel function as Radial Basis Function (RBF) with the user of *e1071* library in *R*. Training models were tuned with 10-fold cross validation to find the optimal parameters and the testing was done on the remaining test data set with the best fitting SVM model.

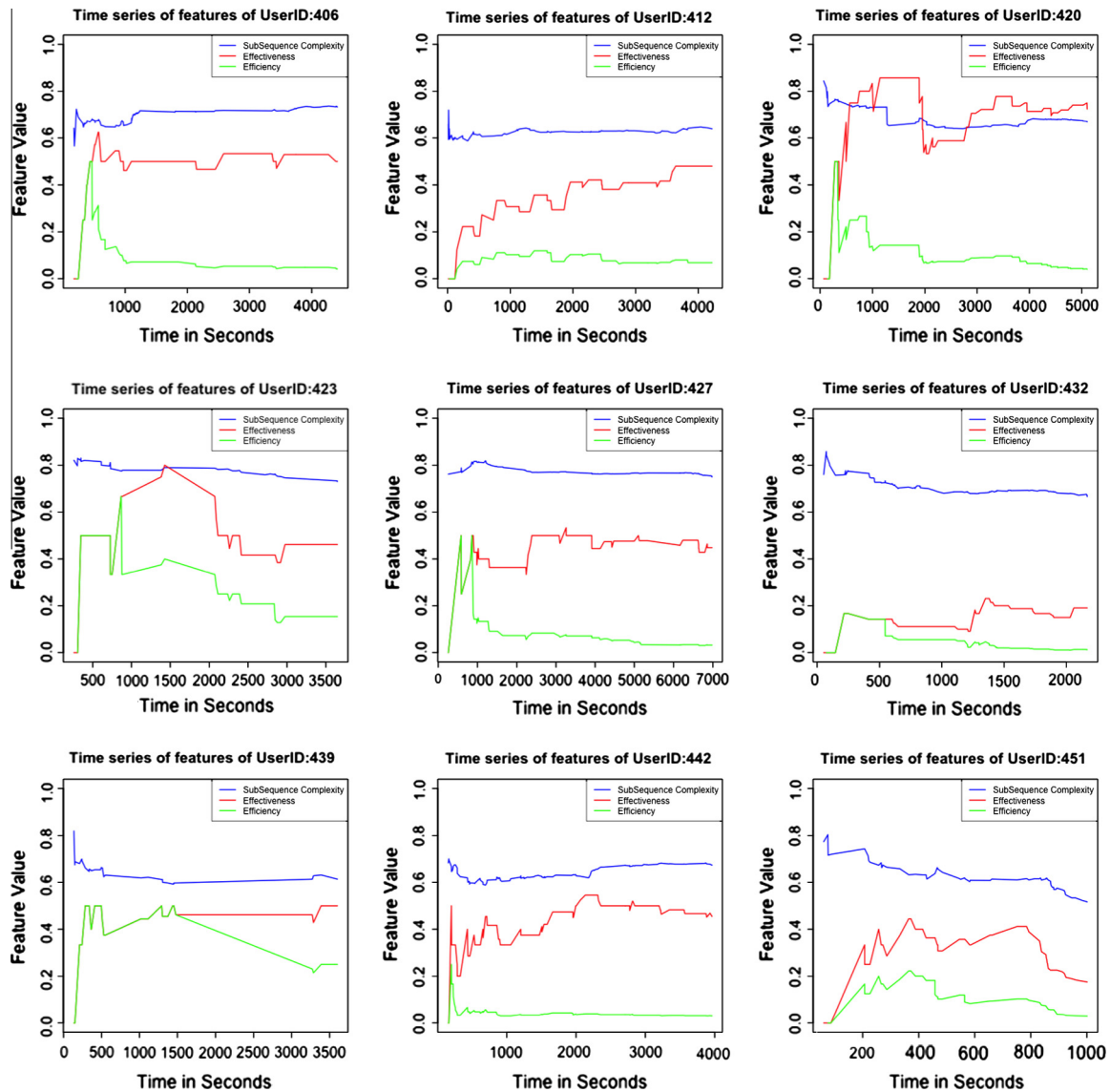


Fig. 6. Time-based implicit features for users for Session 1.

We then constructed the feature matrices, which consist of all aforementioned features for each minute of time t for all the users conducting an exploratory search task. This was converted into a long vector of features, which we then fed as the input to the classification models used. The class labels were generated as 0 or 1 at minute $t + n$ based on the above mentioned criteria and used as the output class labels to be used in the n -steps-ahead prediction model. If a class label at n -steps-ahead was correctly predicted based on the features extracted up to time t from the classification model, it was considered as correctly classified and, if not, as misclassified. For the baseline evaluation, the precision and recall features up to each time t was calculated and the class labels for the classification remained the same as the binary classification mentioned above.

The classification accuracies for the test data for each topic based on the trained SVM models are shown in Table 2. It can be seen that in almost all cases the implicit features based classification led to higher accuracies than the baseline method.

6.2. Simulations of search process recommendations

Having identified an effective way to evaluate search performance that is user-focused and independent of the system itself, we asked if such an evaluation could help in improving performance by process-based recommendation. In other words, in order to evaluate whether users who are predicted to perform sub par in the future based on their current search process could benefit from this analysis to improve their search process, we conducted a few simulations.

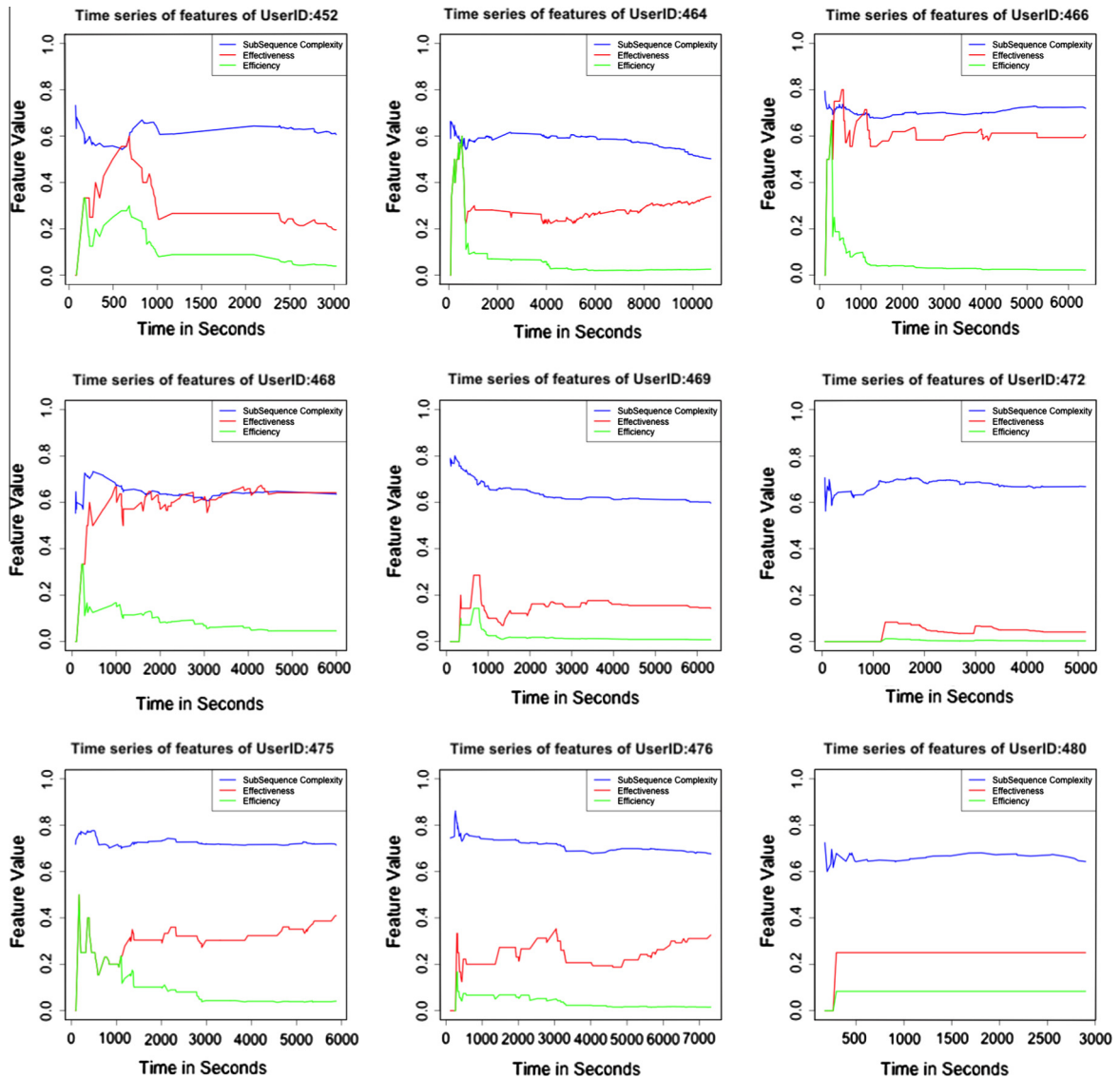


Fig. 7. Time-based implicit features for users for Session 1 cont'd.

Table 2

Classification accuracies at different time and prediction steps. Results with significance at $p < 0.05$ is indicated by *.

Task topics	Time $t = 10$ min $n = 3$ -steps-ahead		Time $t = 20$ min $n = 5$ -steps-ahead		Time $t = 30$ min $n = 10$ -steps-ahead	
	Baseline (%)	Proposed (%)	Baseline (%)	Proposed (%)	Baseline (%)	Proposed (%)
Technology	45.0	68.0*	46.0	72.0*	63.0	77.0*
Health	55.0	72.0*	59.0	78.0*	55.0	75.0*
Art/history	50.0	77.5*	72.0	73.0*	48.0	78.0*
Entertainment	45.0	65.0*	30.0	68.0*	44.0	65.0*

We considered the individual user search processes as a collection of discrete search paths, where each search path is defined as the search process from the time a user issues a query up to the time user issues a very different query (as proposed by White & Huang (2010)). To find difference between two queries, Generalized Levenshtein (edit) Distance (Levenshtein, 1966), which is a commonly used distance metric for measuring the distance between two character sequences, was used. If the Levenshtein (edit) distance between two subsequent queries was greater than a specific threshold, we considered the search process from the former query to the next query as a single search path. It should be noted that

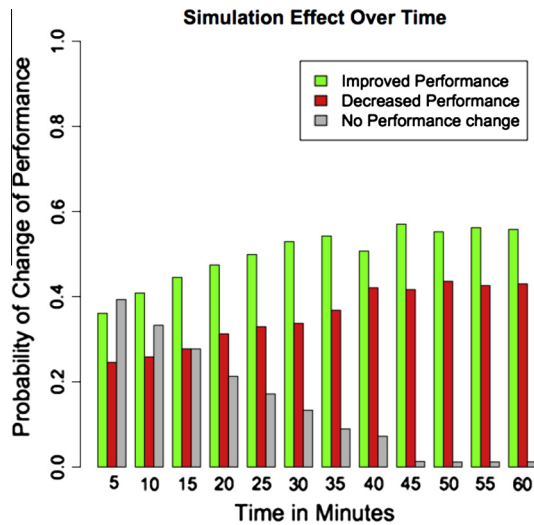


Fig. 8. Probability of change in performance by switching.

Table 3
Performance improvement or decline.

Performance	Mean	Std. dev.
Improvement	0.32	0.14
Decline	0.15	0.08

prior to query similarity calculation, some pre-processing was done to extract the main terms of the query strings by removing stop words (articles, conjunctions, question words, etc) and punctuations.

Algorithm 1. Search Process Recommendation Simulation Algorithm.

Data: Search paths of each user
Result: Simulated search path recommendations for low performers
while *Remaining low performers(l)* **do**
 Find matching high performers with similar first search path;
 for $h=1; h \leq \text{num}(\text{high performers}); h++$ **do**
 if $\text{QueryEditDist}(\text{First search path}(l) - \text{First search path}(h)) \leq \text{Threshold}$
 then
 Switch remaining search path(l) with remaining search path(h);
 else
 No simulation;
 end
 end
end

Following this method, we identified the first search path of each user. Based on the features extracted up to the end of the first search path and the classification model, we learned from the corresponding n -steps-ahead prediction that we could predict whether the user was going to have low/high performance at the end of the session. As the first step (as shown in Algorithm 1), if a user was predicted to have low performance, we then found matching users that had a similar first search path. A threshold of distance 4 was applied based on our observations that when there were changes in the queries due to simple spelling mistakes or subtle query refinements like adding words, this amount was common. For the second step if we found matching users, then we switched the low performing users search process from the end of the first search path with another high performing users remaining search process. We continued doing this until all matching users search processes were injected into the low performing users search process at the end of the first search path. We then analyzed how much performance changed in terms of query effectiveness (QE). This was obtained by evaluating the difference of values for each newly switched search path to the original search path the low performing users were initially following. Once we counted how many of these simulations had an effect in changing the performance level at each time t , we found the probability over

all possible simulations to illustrate how such switching would help, hurt, or not affect the low performing users performance.

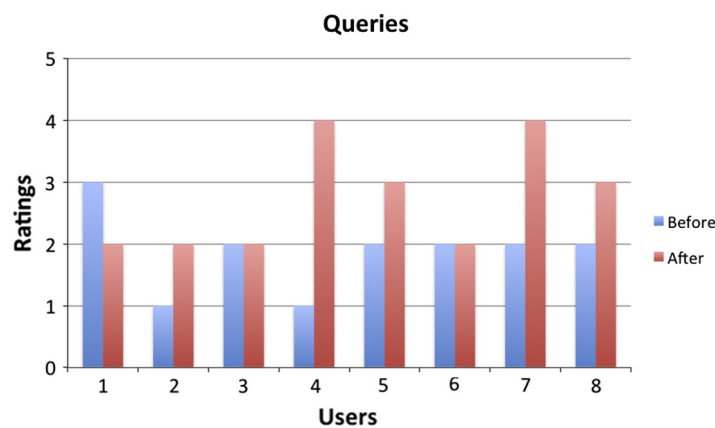
Fig. 8 illustrates the probability of performance change over time for the low performing users by switching to other users search processes who had a similar first search path. They show that switching paths enabled the low performing users to improve their performance mostly during the mid to late part of the search process with probabilities close to 0.5, surpassing the probability of scoring lower than their original performance. Also noticeable is the fact that except for the first few minutes of a session, the chances of no performance change are decreasing drastically. It is encouraging to observe that the users are able to improve their search performance levels by switching to different search paths that would cover more exploratory search aspects such as creativity and knowledge discovery with the proposed search process based recommendations.

The performance improvement or decline at the end of the search process compared to the actual search process in terms of *Query Effectiveness (QE)* is shown in Table 3. It shows that although there were some cases with performance decline, the performance improvement was on average substantially higher than the decline, thus making the search process simulations effective.

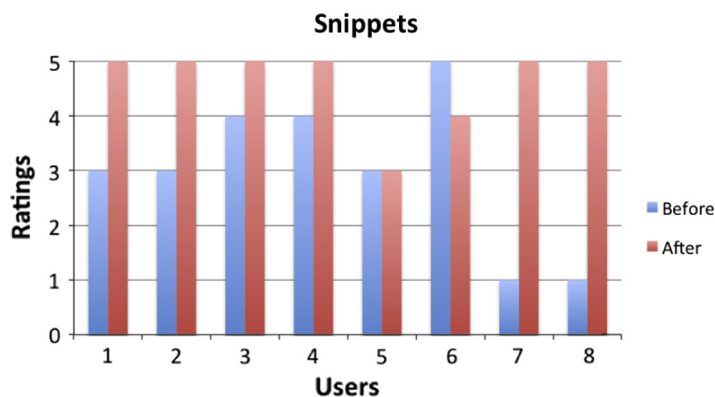
7. Discussion

In addition to the quantitative analysis of simulated search paths (in Section 6.2), we employed an impartial human assessor to evaluate whether the recommendations we provided based on simulations actually made improvements in user's search processes to achieve the task goals. We consulted a professional in Library and Information Science who is qualified to assess and make judgments based on information need and information retrieved.

We provided instructions to rate the complexity of queries and the coverage of snippets collected by each user to the specific topic of exploratory search task. In order to ensure that there was no bias from the assessor, we did not provide information about the user study and did not make the assessor aware of which queries/snippets corresponded to pre/post



(a) Query based assessment



(b) Snippets based assessment

Fig. 9. Human assessment on pre/post recommendation quality on (a) queries (b) snippets.

recommendation. The assessor only had access to the sets of queries and snippets of groups of users working on different topics as well as the task descriptions for each topic. The assessor then provided a qualitative grading on a scale of 1 to 5 (1 being lowest and 5 being highest). The specific instructions provided to the assessor in terms of how to rate the queries and snippets are shown below.

- To rank the queries: Find the complexity of the query terms by counting the number of different unique words (removing stop words such as conjunctions, pronouns, etc) and then give a score of how complex the queries were overall
- To rank the snippets: Rank the users collected snippets based on how well they have collected content to cover the different sub areas mentioned in the task description

Fig. 9 shows the pre/post recommendation assessment for queries and pre/post recommendation assessment for snippets, for 8 users with 2 from each topic area who deemed to be low-performing based on the predictions respectively.

Based on the recommendations 62.5% of the users' query ratings and 75% of the users' snippets ratings has improved. This shows there is a substantial improvement in the user's performance in reaching the task specific goals met by the proposed recommendation strategy.

From the quantitative and human assessment based results, it can be observed that this approach can enable users to improve their search performance by switching to different search paths that would yield a better search experience. This shows that by recommending a search process instead of focusing only on query suggestions and retrieval features, one could facilitate an improved search experience for users conducting exploratory search tasks. This is achieved by evaluating the search process based on implicit features and actions and thereby providing interventions if necessary.

Similar to the Hassan Awadallah, White, Pantel, Dumais, and Wang (2014) the search trail based recommendations provided to the users in our approach improved the overall search performance of the users. Our approach of using implicit measures along aspects of uncertainty, knowledge discovery, creativity and exploration to quantify and assess the performance of user search process when conducting exploratory searches is somewhat similar to White and Huang (2010) who study effectively evaluating search trails evaluated by measuring diversity, coverage, and novelty in exploratory search. Unlike such previous work, our approach predicts the user search performance at different time steps along the search process and thereby, provides recommendations based on other users who deem this process to be successful in improving search performance by switching search trails. The simulations have shown that earlier interventions lead to improved search performance along the search process if the recommendations were taken by the user.

The user study specifically designed to provide a realistic exploratory search experience to the users is effective yet, not without its limitations. We were only able to provide users constrained flexibility to select topics from the 5 topics we used. In a more realistic environment we could ask the users to engage in the topic of their preference but, in this scenario to collect sufficient data from each topic in order to make recommendations we had to impose some constraints. When analyzing the topic selection of users, we noticed that 75% of the users selected a topic that is highly relevant to their declared undergraduate major at the first session. This shows that users are inclined to select an area/topic that they are familiar with when given the choice. We assumed that all participants had similar domain expertise in each topic they selected and did not focus on tuning our simulations for any variations in domain expertise. It would also be interesting to see how this type of study can be expanded by recruiting more participants from varying levels of domain expertise to find the effectiveness and applicability of the simulations in such cases.

Although, this methodology provides substantial evidence of being effective for evaluating and analyzing exploratory search processes and making recommendations, it is not beyond certain limitations. The main limitations of this work is that we assume a priori that a user is conducting an exploratory search task on a selected topic. In other words, the proposed method does not try to predict whether a user is conducting exploratory or non-exploratory searches first before doing the analysis. In addition, we do not focus on evaluating the domain expertise of the users. Our methodology heavily relies on the implicit signals issued by users and all search actions they perform along the search process, thus if there is significant noise whether certain users are doing irrelevant searches and not focusing on the task, we have to carefully look at the data and remove such noise. This means that the data relies significantly on the level of attention and focus a user places on conducting a realistic search task.

8. Conclusion

In the work presented here, we described our attempt to analyze user search behavior at the search process level using implicit features for exploratory search tasks. Most related work so far has focused on *what* users found as artifacts along the search process rather than *how* they found it or *why* they performed such actions to locate information. Therefore, in this paper we tried to understand the underlying search process by using the user search actions and implicit features to develop a framework to analyze exploratory search using a quantitative approach.

We proposed a set of implicit features that could be extracted easily from the transactional log data to capture the underlying search behavior of users conducting exploratory search tasks along the aspects of exploration, knowledge discovery and creativity. We were also able to get insights into the process in which they conduct the task by looking at the search actions sequences and determining their complexity. Once we identified the behavior patterns, we could accurately predict

the search performance level of users few time steps ahead in the process, thus addressing the first research question (RQ a) specified in the Section 1.

In addressing our second research question (RQ b), we proposed a method to assist low performing users in improving their search performance by switching search processes. The search process simulations provide verification for our approach in that by realizing early during the search process whether a user is going to perform well or not, one could recommend better search processes for that user. We showed the effect of simulated recommendation, incorporating both quantitative and qualitative approaches using probability of change in performance and human assessment. Therefore, the use of the proposed method in anticipating n -steps-ahead predictions of a users specific performance level could have potential applications with linkage to recommendation of search processes to users. These could then make useful suggestions to a user at the most appropriate time during an ongoing search process to enhance search performance and thereby gain success to reach the task goals.

Acknowledgements

We would like to thank Matthew Mitsui for providing assistance in designing, developing and conducting the user study. We are also thankful to Danielle Reay for proof-reading the article. We are grateful to Daniel Russell from Google for providing advice and guidance in designing the user study. We express our thanks to all the participants who took part in this study and all those who helped us in pilot testing and conducting the user study in numerous ways.

We would also like to express our sincere gratitude to the reviewers for their valuable comments that helped in strengthening this article.

The work reported here is supported by the United States Institute of Museum and Library Services (IMLS) Early Career Development Grant #RE-04-12-0105-12.

References

- Al Bawab, Z., Mills, G. H., & Crespo, J.-F. (2012). Finding trending local topics in search queries for personalization of a recommendation system. In *Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining (KDD '12)* (pp. 397–405). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/2339530.2339594>.
- Azzopardi, L. (2014). Modelling interaction with economic models of search. In *Proceedings of the 37th international ACM SIGIR conference on research & development in information retrieval (SIGIR '14)* (pp. 3–12). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/2600428.2609574>.
- Bawden, D. (1986). Information systems and the stimulation of creativity. *Journal of Information Science*, 12(5), 203–216. URL <http://dx.doi.org/10.1177/016555158601200501>.
- Bennett, P. N., Radlinski, F., White, R. W., & Yilmaz, E. (2011). Inferring and using location metadata to personalize web search. In *Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval (SIGIR '11)* (pp. 135–144). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/2009916.2009938>.
- Bennett, P. N., White, R. W., Chu, W., Dumais, S. T., Bailey, P., Borisyuk, F., et al (2012). Modeling the impact of short- and long-term behavior on search personalization. In *Proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval (SIGIR '12)* (pp. 185–194). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/2348283.2348312>.
- Borgman, C. L., Hirsh, S. G., & Hiller, J. (1996). Rethinking online monitoring methods for information retrieval systems: From search product to search process. *American Society for Information Science*, 47(7), 568–583. doi: 10.1002/(SICI)1097-4571(199607)47:7<568::AID-ASIS>3.0.CO;2-S.
- Borlund, P., & Ingwersen, P. (2006). The application of work tasks in connection with the evaluation of interactive information retrieval systems: Empirical results. In *Proceedings of the 1999 international conference on final mira (MIRA'99)*. Swinton, UK: British Computer Society, pp. 1–1. <<http://dl.acm.org/citation.cfm?id=2228065.2228066>>.
- Cole, M. J., Hendahewa, C., Belkin, N. J., & Shah, C. (2014). Discrimination between tasks with user activity patterns during information search. In *Proceedings of the 37th international ACM SIGIR conference on research & development in information retrieval (SIGIR '14)* (pp. 567–576). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/2600428.2609591>.
- Foster, A., & Ford, N. (2003). Serendipity and information seeking: An empirical study. *Journal of Documentation*, 59(3), 321–340. <http://www.emeraldinsight.com/doi/abs/10.1108/00220410310472518>.
- Fox, S., Karnawat, K., Mydland, M., Dumais, S., & White, T. (2005). Evaluating implicit measures to improve web search. *ACM Transaction on Information Systems*, 23(2), 147–168. <http://doi.acm.org/10.1145/1059981.1059982>.
- Gabadinho, A., Ritschard, G., Studer, M., & Muller, N. S. (2010). Indice de complexité pour le tri et la comparaison de séquences catégorielles. *Revue des nouvelles technologies de l'information (RNTI) E-19*, 61–66.
- Hassan Awadallah, A., White, R. W., Pantel, P., Dumais, S. T., & Wang, Y.-M. (2014). Supporting complex search tasks. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management (CIKM '14)* (pp. 829–838). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/2661829.2661912>.
- Hendahewa, C., & Shah, C. (2013). Segmental analysis and evaluation of user focused search process. In *2013 12th International conference on machine learning and applications (ICMLA)* (Vol. 1, pp. 291–294).
- Kuhlthau, C. (2004). Seeking meaning: A process approach to library and information services. Information management, policy, and services. Libraries unlimited. <<http://books.google.com/books?id=feDgAAAAAAAJ>>.
- Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, 10, 707.
- Liu, J., & Belkin, N. J. (2012). Searching vs. writing: Factors affecting information use task performance. *Proceedings of the American Society for Information Science and Technology*, 49, 1–10.
- Marchionini, G. (2006). Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4), 41–46. URL <http://doi.acm.org/10.1145/1121949.1121979>.
- Matthijs, N., & Radlinski, F. (2011). Personalizing web search using long term browsing history. In *Proceedings of the fourth ACM international conference on Web search and data mining (WSDM '11)* (pp. 25–34). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/1935826.1935840>.
- Pirolli, P., & Card, S. K. (1999). Information foraging. *Psychological Review*, 106, 643–675.
- Qvarfordt, P., Golovchinsky, G., Dunnigan, T., & Agapie, E. (2013). Looking ahead: Query preview in exploratory search. In *Proceedings of the 36th international ACM SIGIR conference on research and development in information retrieval (SIGIR '13)* (pp. 243–252). New York, NY, USA: ACM. URL <http://doi.acm.org/10.1145/2484028.2484084>.

- Ruthven, I. (2003). Re-examining the potential effectiveness of interactive query expansion. In *Proceedings of the 26th annual international ACM SIGIR conference on research and development in informaion retrieval (SIGIR '03)* (pp. 213–220). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/860435.860475>.
- Saracevic, T. (1997). The stratified model of information retrieval interaction: Extension and applications. *Proceedings of the Annual Meeting–American Society for Information Science*, 34, 313–327.
- schraefel, m. c., Wilson, M. L., Russell, A., & Smith, D. A. (2006). mspace: Improving information access to multimedia domains with multimodal exploratory search. *Communication of the ACM*, 49(4), 47–49. <<http://eprints.soton.ac.uk/262376/>>.
- Shah, C., & González-Ibáñez, R. (2011). Evaluating the synergic effect of collaboration in information seeking. In *Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval (SIGIR '11)* (pp. 913–922). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/2009916.2010038>.
- Shah, C., & González-Ibáñez, R. (2012). Spatial context in collaborative information seeking. *Journal of Information Science*, 38(4), 333–349. <http://dx.doi.org/10.1177/0165551512438356>.
- Shah, C., Hendahewa, C., & González-Ibáñez, R. (2015). Rain or shine? forecasting search process performance in exploratory search tasks. *Journal of the Association for Information Science and Technology*. <http://dx.doi.org/10.1002/asi.23484>.
- Singh, R., Hsu, Y.-W., & Moon, N. (2013). Multiple perspective interactive search: A paradigm for exploratory search and information retrieval on the web. *Multimedia Tools and Applications*, 62(2), 507–543. <http://dx.doi.org/10.1007/s11042-011-0910-2>.
- Smucker, M. D., & Clarke, C. L. (2012). Time-based calibration of effectiveness measures. In *Proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval (SIGIR '12)* (pp. 95–104). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/2348283.2348300>.
- Spink, A. H. (2002). A user-centered approach to evaluating human interaction with web search engines: An exploratory study. *Information Processing and Management*, 38(3), 401–426. <<http://eprints.qut.edu.au/4993/>>.
- Sugiyama, K., Hatano, K., & Yoshikawa, M. (2004). Adaptive web search based on user profile constructed without any effort from users. In *Proceedings of the 13th international conference on World Wide Web (WWW '04)* (pp. 675–684). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/988672.988764>.
- Vakkari, P. (2001). A theory of the task-based information retrieval process: A summary and generalization of a longitudinal study. *Journal of Documentation*, 44–60.
- Vakkari, P., & Huuskonen, S. (2012). Search effort degrades search output but improves task outcome. *American Society for Information Science*, 63, 657–670.
- White, R. W., & Drucker, S. M. (2007). Investigating behavioral variability in web search. In *Proceedings of the 16th international conference on World Wide Web (WWW '07)* (pp. 21–30). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/1242572.1242576>.
- White, R. W., & Huang, J. (2010). Assessing the scenic route: Measuring the value of search trails in web logs. In *Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval (SIGIR '10)* (pp. 587–594). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/1835449.1835548>.
- White, R. W., Muresan, G., & Marchionini, G. (2006). Report on acm sigir 2006 workshop on evaluating exploratory search systems. *SIGIR Forum*, 40(2), 52–60. <http://doi.acm.org/10.1145/1189702.1189711>.
- White, R. W., & Roth, R. A. (2009). Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1), 1–98.
- Wildemuth, B. M. (2004). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, 55(3), 246–258. <http://dx.doi.org/10.1002/asi.10367>.
- Wildemuth, B. M., & Freund, L. (2012). Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the symposium on human-computer interaction and information retrieval (HCIR '12)* (pp. 4:1–4:10). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/2391224.2391228>.
- Wilson, M. L., schraefel, M. C., & White, R. W. (2009). Evaluating advanced search interfaces using established information-seeking models. *Journal of the American Society for Information Science and Technology*, 60(7), 1407–1422. <http://dx.doi.org/10.1002/asi.v60:7>.
- Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '03)* (pp. 401–408). New York, NY, USA: ACM. <http://doi.acm.org/10.1145/642611.642681>.
- Zhang, J., & Marchionini, G. (2005). Evaluation and evolution of a browse and search interface: Relation browser. In *Proceedings of the 2005 national conference on digital government research* (pp. 179–188).