Analyzing User Behavior Patterns in Adaptive Exploratory Search Systems with LifeFlow

Jae-wook Ahn, Krist Wongsuphasawat
Human-Computer Interaction Lab
University of Maryland
{jahn,kristw}@cs.umd.edu

Peter Brusilovsky
School of Information Sciences
University of Pittsburgh
peterb@pitt.edu

ABSTRACT

Adaptive exploratory search is a method that can provide user-centered personalized search results by incorporating interactive user interfaces. Analyzing the user behavior patterns of these systems can be complicated when they support transparent and controllable open user models. This paper suggests to use a visualization tool to address the problem, as a complement to the typical statistical analysis. By adopting an event sequence visualization tool called LifeFlow, we were able to easily find out user interesting behavior patterns, especially regarding the open user model exploration.

Keywords

Personalized Search, Open User Model, Exploratory Search, Human-Computer Interaction, Visualization

1. INTRODUCTION

Adaptive exploratory search systems (ESS) can provide efficient user-centered personalized search results by incorporating interactive user interfaces. The core component of the adaptive systems is called a user model, where the user's search tasks, contexts, and interests are stored, so that the systems can adapt to those specific tasks or contexts [7]. In conventional adaptive search approaches, the user models are hidden from the users. They are *black boxes* and the users cannot estimate easily what is going on inside. Therefore, they cannot expect what personalized results will be returned from the black box and have very limited capability to correct the system's unexpected erroneous behaviors. Several approaches including our own [3] tried to solve this problem by transparently reveal the user model contents and provide the ability to directly control the user models.

The open user model also adds exploratory nature to adaptive search systems. By combining the search algorithms and interactive user interfaces, ESS can let users promote their knowledge and expressive power about their problems and learn about the search space [6]. Just like this idea, open user models clearly show users what their personal mental models on the search tasks are and how the search processes evolve. Over the years, we have tried this line of ideas incrementally: keyword-cloud style user model viewer/editor for news filtering [3], user query/user model fusion controller [4], and a completely visual user model approach [1]. They helped users visually understand how their user model contents change dynamically, directly manipulate the user models by adding or removing keywords, and control the relative importance of the user models compared to user queries. This higher transparency and controllability is one of the greatest virtues of adaptive ESS. However, it increased the complexity when we attempted to analyze the behaviors of its users.

We need to find out interesting patterns or sequences of user activities, which are much more diverse compared to traditional look-up search activities. If we can learn the behavioral patterns of the users of adaptive ESS, we will be able to provide more efficient systems. Typical statistical log file analysis may not meet this need because it could be harder to find out unexpected and novel patterns, which we expect to discover. Therefore, we used an event sequence visualization tool called LifeFlow [9]. Using the interactive search log visualization method, we could more quickly find out interesting patterns or anomalies as in [8].

The goals of this study can be stated in two folds. First, we attempted to analyze the search log of ESS and find out interesting user behavior patterns in order to support the design of future adaptive ESS as in [5] using a visualization tool. We could compile a list of frequent user action sequences and find out the characteristics of highly visible and interactive adaptive ESS. The second goal is more focused on the open user model manipulation behaviors. We could observe what sequences the users had taken when they accessed the core functionality of the adaptive ESS.

2. ADAPTIVE EXPLORATORY SEARCH SYSTEMS: TASKSIEVE AND ADAPTIVE VIBE

We highlight two adaptive ESS with open user models called TaskSieve [4] and Adaptive VIBE [2, 1] (Figure 1). TaskSieve uses typical ranked lists but it has a component for adjusting the importance of user model and user queries. Users can explore different search results by switching three sets of weight configurations of user models versus user queries (1:0, 0.5:0.5, and 0:1) while monitoring their user model contents. Adaptive VIBE extends the open user model using a reference point based interactive visualization called VIBE (Visual Information Browsing Environment). It defines the user query and the user model (as the list of keywords) as reference points (Point of Interest or POIs) and interactively places the search result documents according to their similarities to each POI. Documents are placed closer to more similar POIs and users drag the POIs in order to visually explore the search space. By adopting the user model as POIs, Adaptive VIBE could seamlessly implement visual and editable user models. We have shown that the two systems could provide better performance than non-adaptive search

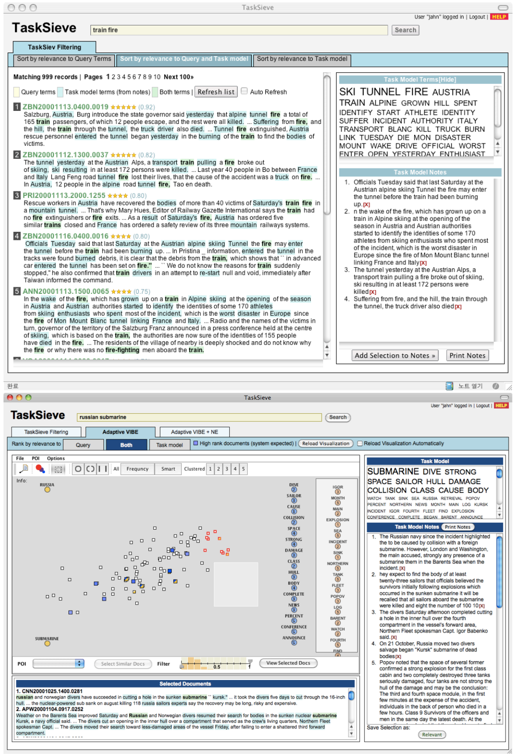

Figure 1: TaskSieve (above) and Adaptive VIBE (below)

systems without open user models in term of search precision and diversity [4, 1, 2]. Both systems have log facilities and the user activities in the log was used for the analysis. These actions are organized by higher-level categories in Table 1 and 2.

3. ANALYSIS METHOD

We analyzed user activity sequences from the log files of the two systems. The log data was extracted from a user study including 30 users in 60 sessions (20 minutes per session). For each session, the systems were loaded with news articles from the Topic Detection and Tracking (TDT4) test collection and the participants of the user study were asked to finish search tasks. Our goal of this study was to discover two user behavior patterns from the log data.

1. **Switch of activities** – Typical look up search systems will have simple activity switch patterns such as *query* > *examine list* > *examine documents*. We expected to see different patterns from the open user model based ESS. Users were supported with more features to explore the search space and view/control their mental models. We were interested to discover how the users really exploited those features.

2. **User activity sequence patterns** – It would be useful to have a list of user activity patterns about how users actually used the features while trying to solve specific sub-tasks. This knowledge will help improving the future ESS systems with open user models.

In order to accomplish these goals, we needed a method that could easily analyze complicated log files filled with

Table 1: Adaptive VIBE User Actions

Category	Action
Login/out	login, logout
Search	search
Overview	reset visualization
Find subset	select similar documents from a POI, marquee selection, double slider filtering, show similar documents or POIs, view selected documents, view auto-selected documents
Examine doc	view by mouse-over, select doc, open doc
POI activities	reset poi, select poi, select query poi, move POI, move query POI
Manipulate UM (User Model)	change UM & query weight, select UM poi, move UM poi
Update UM contents	save note, remove note

Table 2: TaskSieve User Actions

Category	Action
Login/out	login, logout
Search	search
Overview	search response, navigate page
Examine doc	open doc
Manipulate UM	change UM & query weight
Update UM contents	save note, remove note

the diverse user activities. We adopted a visualization tool called LifeFlow [9] that supported analyzing event sequences. It was originally designed for electronic health records analysis but keeps discovering more applications (e.g. transportation). It summarizes the sequence patterns and makes possible the comparison of the patterns. Among the rich set of methods it supports, we found the followings were relevant to solve the above goals of this study.

1. **Overview of temporal sequences** – The main feature of LifeFlow is to show the list of sequences and then sort them by the frequency. Using a tree structure, it shows the hierarchy of the sequences and their frequencies that reveal what sequence of activities are most dominant.

2. **Align by a specific sequence** – We can select a specific event and then align other activities before/after the selected event. They are ordered and visualized by the frequencies again, so that we can see which activities were more frequent before and after a specific event.

We generated two datasets from the log data as inputs to the two analysis above: (1) list all actions (per search session of a user) in a row for activity switch analysis and (2) pairs (i.e. bigrams) of actions to observe the user activity sequence patterns.

4. RESULT ANALYSIS

4.1 Activity Switch Analysis

Figure 2 compares the overviews of the search activity sequences of TaskSieve and Adaptive VIBE. Each row (horizontal bar) represents a 20 minute search session and the colored blocks mean specific actions taken by the participants.

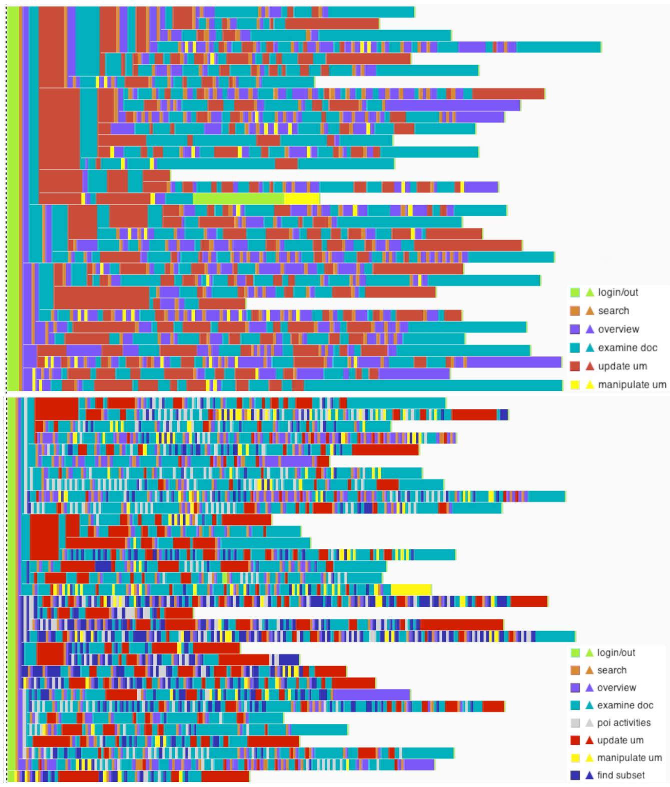

Figure 2: Visual Overview of User Behaviors in TaskSieve (above) and Adaptive VIBE (below)

The width of the colored blocks corresponds to the duration of the action. We used the high-level action categories (Table 1 and 2) instead of individual actions in order to more clearly understand the meaning of them. The two most frequent activities that occupied the largest portion of the screens were **Update UM** (*red*) and **Examine Doc** (*dark green*). It had been expected because the users would spend significant amount of time reading the snippets or fulltexts of the documents (**Examine Doc**) and edit their notebooks. In two systems, notebook editing action automatically leads to updating the user model (**Update UM**). What was more interesting was the density of the exploratory activities. In the visualization-based system (Adaptive VIBE), the distribution of different actions were denser than in the text-based system (TaskSieve). The participants switched more frequently by spending less time (smaller width colored blocks) per action. This observation reflects that they fully exploited the richer feature set provided by Adaptive VIBE. At the same time, the high action switches included a lot of exploratory actions. We could assume that they were able to take advantage of the flexibility of the visualization system and performed more exploratory behaviors.

The second observation is the user model manipulation. Users could directly manipulate the user model weights or drag the user model keywords (POIs), in order to better reorganize the visualization of retrieved documents, learn about and control the effects of their user models. In Adaptive VIBE, these activities (*yellow*) were more frequent and prevalent, which suggests that the users more actively used the visual user model manipulation feature of the system.

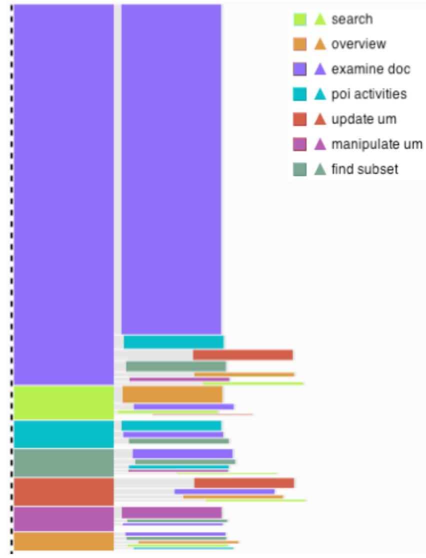

Figure 3: Pairwise distribution of user activities

4.2 User Activity Pattern Analysis

4.2.1 *List of frequent activity patterns*

The second part of the analysis was to find out the exact user behavior patterns. It will help to design more efficient adaptive exploratory search systems if we could learn how the users really used them. In this paper, we limited this analysis to Adaptive VIBE only, because it provided more diverse adaptive exploration features than TaskSieve. Figure 3 shows the distribution of pairwise user actions of Adaptive VIBE. They are bigrams of sequential user actions across the whole sessions. This time, vertical height means the frequency of the pairwise action patterns and the width of the colored blocks is the mean duration time.

LifeFlow shows the pairs in the tree structure – the first action of the pair as the first level and the second action as the second level in it. Therefore, we could analyze the pairs hierarchically – first find out the most frequent actions and then the following frequent actions. For example, the most frequent action is **Examine Doc** (*blue*). In the second level, the most frequent action was **Examine Doc** again but the second and the third were **POI Activities** and **Update UM**, which means that the users moved/selected POIs (not user model POIs in this case) or updated their user model contents by adding or removing texts to/from the notebook. This hierarchical tree could make the structured analysis of the pairs easier than simple frequency counting, by following down the branches of dominant activities in the first level.

Table 3 summarizes the list of important pairs discovered from this analysis. Overall, the frequencies of the first actions of the pairs (bar height) were almost equivalent except **Examine Doc**. We thus included every seven activity for the first level and then counted the top three leaves in the next second level. The findings from the table are as follows: (1) A lot of exploratory actions – examine the overview, control the visualization to change/re-interpret the big picture (including the *user model update/manipulation*), zoom and find subset – were much more prevalent compared to simple look-up search action, (2) There are a lot of repeating sequences – (examine doc, examine doc), (poi activities, poi

Table 3: List of dominant user activity pairs

1st activity	2nd activity	Frequency
examine doc	examine doc	6,392
examine doc	poi activities	254
examine doc	update um	204
search	overview	330
search	examine doc	108
search	search	48
poi activities	poi activities	192
poi activities	examine doc	107
poi activities	find subset	102
find subset	examine doc	178
find subset	find subset	93
find subset	poi activities	62
update um	update um	193
update um	examine doc	97
update um	overview	62
manipulate um	**manipulate um**	**222**
manipulate um	**find subset**	**53**
manipulate um	**examine doc**	**37**
overview	examine doc	63
overview	find subset	58
overview	overview	46

activities), (update um, update um), (manipulate um, manipulate um), (3) User model exploration was used almost as frequently as other actions, (4) This table can be used as a list of possible user actions for similar adaptive ESS.

4.2.2 User model exploration activities

The next step was to look deeper into the user activities regarding how the users explored and controlled their open user models, which were the core module of the adaptive ESS. Adaptive VIBE implemented the visual user models as keyword POIs and let users drag them around the screen, select them for supporting further actions (e.g. selecting documents similar to the selected POI), disable some of them temporarily, etc. From the analysis above, these user model manipulation activities were conducted as significantly as the others – the three most frequent actions *just after* **Manipulate UM** were **Manipulate UM** (*purple*), **Examine Doc** (*blue*), and **Find Subset** (*dark green*) (Table 3). This suggests that the user model manipulation actions were done repeatedly and the users could narrow down their search targets as a results of user model exploration. Now we can focus more on the user model exploration by aligning the sequences by the user model manipulation actions. In Figure 4, the right half after the vertical dashed line is the pairs where **Manipulate UM** was preceding the second action. The left half is vice versa. From this left half of the *align* view of LifeFlow, we could observe that the three most frequent actions *after* **Manipulate UM** were still the most frequent actions *before* **Manipulate UM**. It leads us to conclude that the user model manipulations were in the chain of the three actions repeatedly and the user model manipulation task needs to be considered as a set with those *friend* actions. If we look back to the overview (Figure 2 below), we can confirm the repeating patterns of the manipulate UM (*yellow*), examine doc (*dark green*), find subset (*dark blue*) are adjacent with each other.

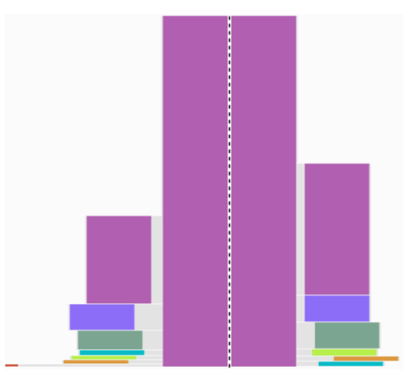

Figure 4: Pairwise user activities aligned by "User Model Manipulate" action, showing the activities *before* and *after* them (same copy key with Figure 3).

5. CONCLUSIONS

This paper provides a preliminary analysis of user behavior patterns in adaptive exploration search systems. We analyzed the log data of a user study that included various user actions using an event sequence visualization tool called LifeFlow. From the analysis, we could find that the user exploration actions were switching more frequently using the visualization-based adaptive ESS system. At the same time, the user model exploration features were used as frequently as other features and we could find patterns about how the user model explorations were done combined with other exploratory behaviors. We also generated a list of user action patterns of the adaptive ESS systems for supporting future system design of the same kind. Our future plan includes deeper analysis of other activities closely related to user model exploration, such as finding subset, query exploration, and user model updates.

6. REFERENCES

[1] J. Ahn. *Adaptive Visualization for Focused Personalized Information Retrieval*. PhD thesis, University of Pittsburghty, September 2010.

[2] J. Ahn and P. Brusilovsky. Adaptive visualization of search results: Bringing user models to visual analytics. *Information Visualization*, 8(3):167–179, 2009.

[3] J. Ahn, P. Brusilovsky, J. Grady, D. He, and S. Y. Syn. Open user profiles for adaptive news systems: help or harm? In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 11–20, New York, NY, USA, 2007. ACM Press.

[4] J. Ahn, P. Brusilovsky, D. He, J. Grady, and Q. Li. Personalized web exploration with task models. In *WWW '08: Proceedings of the 17th international conference on World Wide Web*, pages 1–10, New York, NY, USA, 2008. ACM.

[5] B. Kules, R. Capra, and J. Ryan. Transaction log analysis of user actions in a faceted library catalog interface. In *Fourth Workshop on Human-Computer Interaction and Information Retrieval (HCIR 2010)*, 2010.

[6] G. Marchionini. Exploratory search: from finding to understanding. *Communications of ACM*, 49(4):41–46, April 2006.

[7] A. Micarelli, F. Gasparetti, F. Sciarrone, and S. Gauch. The adaptive web. chapter Personalized search on the world wide web, pages 195–230. Springer-Verlag, Berlin, Heidelberg, 2007.

[8] R. White, A. Kapoor, and S. Dumais. Modeling Long-Term Search Engine Usage. *User Modeling, Adaptation, and Personalization*, pages 28–39, 2010.

[9] K. Wongsuphasawat, J. A. Guerra Gómez, C. Plaisant, T. D. Wang, M. Taieb-Maimon, and B. Shneiderman. Lifeflow: visualizing an overview of event sequences. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, CHI '11, pages 1747–1756, New York, NY, USA, 2011. ACM.