Interaction Model to Predict Subjective–Specificity of Search Results

Kumaripaba Athukorala [‡], Antti Oulasvirta [†], Dorota Glowacka [‡], Jilles Vreeken [†], Giulio Jacucci [‡]

[‡] Helsinki Institute for Information Technology HIIT
[‡] Department of Computer Science, University of Helsinki
[†] Max Planck Institute for Informatics and Saarland University

Abstract. Exploratory search is becoming more common as the web is used more increasingly as a medium for learning and discovery. Compared to traditional known-item search, exploratory search is more challenging and difficult to support because it initiates with poorly defined search goals, while the user knowledge and information needs constantly change throughout the search process. Modeling the user behavior in exploratory search is a hard problem to solve. In spite of a large amount of research on personalization, little attention has been devoted to personalization in the context of exploratory search taking into account the evolving information needs of the user. We propose a formal model—motivated by Information Foraging Theory—for predicting specificity of search results with respect to the evolving knowledge and information needs of the user in exploratory search.

Keywords: User Modeling, Exploratory Search, Scientific Information-Seeking, Search Result Specificity, Information Foraging Theory

1 Introduction

Nowadays the web is used increasingly more as a source for learning and exploratory discovery [1]. Most of the existing information retrieval (IR) systems provide adequate support for well defined information needs [2]. However, there is still room for improvement for current IR systems to support users in situations where the search goal is ill-defined and changes as the search progresses, users lack the knowledge to formulate queries that express their information needs clearly, and users struggle in complex information spaces [2]. Researchers from diverse communities such as IR, machine learning or human computer interaction (HCI) have been working on designing search engines, user interfaces and user models to better support this kind of searches commonly referred to as "exploratory search."

Over the last decade many techniques have been proposed to provide better support for exploratory search, such as results clustering [3], relevance feedback [4], faceted search [5], as well as novel visualizations to support the exploration of unfamiliar information spaces [6]. However, evidence from user studies suggests that results clustering, faceted search, and relevance feedback based methods are rarely used due to the high cognitive overload of selecting relevant results and providing feedback for a large number of items [5, 4]. In response, a number of new techniques were designed to visualize search results and capture user feedback. Some of them include rich user interfaces combined with learning algorithms to support users to comprehend the search results [6], and visualization and summaries of results [7]. All these solutions are giving users more control, however, they fail to take the moment-by-moment information-needs of the user into consideration [8].

Exploratory search involves many different phases. For example, users begin exploring an unfamiliar information space by formulating imprecise queries because they lack knowledge to express their information needs [1]. Then, through several successive iterations of exploring the retrieved information and reformulating queries, the scope of the information need might narrow down [2]. This iterative and evolving nature of exploratory search makes it difficult for IR systems to identify the constantly changing information needs of the user and different phases of exploration. This is where user modeling can greatly improve existing approaches to exploratory search.

There already exist some research on modeling query formulation and interaction strategies to predict the user knowledge and information needs to personalize search results. For example, [9] presents a model which predicts the user knowledge from eye movement patterns. Even though such a model is useful in identifying domain novices and experts, it cannot predict how the information need of a user changes in a search session. Research on the correlation between the length of search queries and specificity of the user information needs [10] suggest that the length of a search query is positively correlated with the specificity of the user information-need. Such a model is useful to predict whether the information needs are too specific or broad. However, users may express very specific information needs with narrower queries having specific keywords at all query lengths and such a correlation based model cannot predict this scenario.

Systems that suggest or expand queries, provide interactive keyword visualizations, cluster results to better support exploratory search need to "know" whether the results generated from such suggestions are too broad/narrow for the information needs of the user. Hence, in exploratory search it is important to predict whether the search engine result pages (SERPs) are too narrow/broad for the evolving information needs of the user. There exist user models for navigational/transactional searches, that predict user satisfaction and relevance of SERPs from behavioral signals such as search result clicks, query refinements, gaze distribution, and dwell times [11]. Even though they provide implicit relevance feedback to the search engine, they do not predict whether the future search results should be narrower/broader. To the best of our knowledge, there has been no work on predicting the specificity (whether the results are broad or narrow) of SERPs to user knowledge and information needs.

One way to address this problem is by understanding user behaviors with queries that retrieve SERPs with varying subjective specificity in exploratory searching, which, in turn, will allow us to build a user model to predict whether given SERPs are too broad or too narrow for the current information needs of the user.

In this short paper, we will briefly discuss the following issues:

- Relationship between result click rate and specificity of SERPs to user information needs;
- A formal model to predict the specificity of SERPs to user information needs at three levels: broad, intermediate, specific;
- Empirical validation of the model.

2 Interaction Model Overview

We designed a model of user interaction by combining insights from research into exploratory search and Information Foraging Theory (IFT) [12]. According to IFT, information gain can be modeled as a linear function of time when the results are not ordered by relevance to the query. Further, IFT states that this information gain function will qualitatively shift towards a diminishing returns curve if results are ordered by relevance to the query. The gradient of this information gain function can be further improved by introducing new interface elements, such as result clustering. Hence, IFT shows how information gain is affected by the user interface or system changes.

Our research is motivated by this model. We use the term *subjective–specificity* to refer to the specificity of SERPs to the user knowledge and information need. By information need we mean the type of information that the user is actually interested in. If we keep the user interface constant, the information gain function should change according to the subjective–specificity of search results. We define subjective–specificity at three levels: broad, intermediate, and narrow. If a user issues search queries that retrieve SERPs covering many diverse topics, then we refer to them as having *broad* subjective–specificity. For example, consider an undergraduate who has just begun to follow a course in data mining issuing "data mining" as her first query to explore this domain. However, for a broad subject like data mining users would benefit from visualizations that provide an overview of the information space [13]. If a user issues search queries retrieving SERPs that are referring to a sub-topic in this domain, such as "pattern mining" under the subject data mining, then we refer to it as having *intermediate* level of subjective–specificity. In exploratory search users passively obtain cues about new keywords and repetitively reformulate queries based on these cues [2]. However, the SERPs that the user retrieves with these new keywords might be too specific for the users information need. For example, consider the same user issuing the query "subgroup discovery" based on the keywords the search engine suggested or s/he has noticed in the previous SERPs. The SERPs for this query might cover a very narrow topic, containing technical details that are less comprehensible for a novice in that area. We refer to such SERPs as having *narrow* subjective–specificity. Generally, for such a narrow search query, a novice user would benefit from more introductory material about the topic such as Wikipedia articles, book chapters, and literature reviews as well as more guided support through the specific information space [14]. The key idea is that the same search result can have very different information content for a user depending on how well it matches their current information needs. If a search engine can predict the subjective–specificity of SERPs (i.e. as broad or narrow) according to such changes in user's information need then it would be very useful to provide more related and personalized results to the user.

Our model captures how information gain in exploratory search is affected by this subjective–specificity. We define the information gain function as given in Equation 1:

$$g(n) = \lambda \ln(n) - \alpha \tag{1}$$

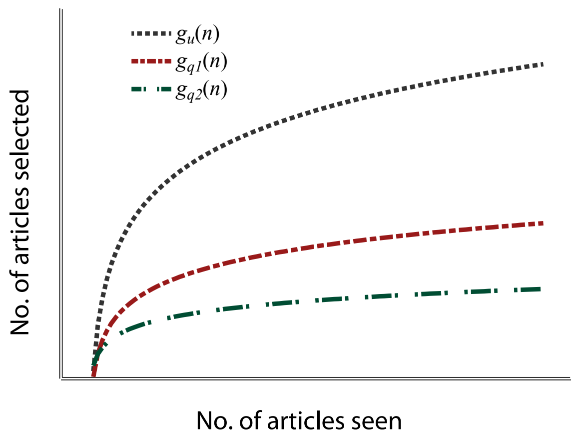

Fig. 1. Hypothetical example of information gain as a function of the number of articles Seen–Selected. $g_u(n)$ is the user-specific effective information gain function. $g_{q1}(n)$ and $g_{q2}(n)$ show how the gradient of the Seen–Selected graph reduces when the subjective–specificity of SERPs is higher than the user's information need.

Here, we define information gain as the number of results *selected* by the user. We refer to the action of clicking a search result in SERPs as selecting. We express the information gain (g) curve of a user as a function of number of result items from SERPs seen by the user (n). Following IFT, we expect this gain function to take the shape of a diminishing returns curve as shown in Figure 1. We refer to this graph as the *Seen–Selected* curve. In this gain function λ determines the slope of the information gain curve. We expect the gradient, λ, to decrease if the SERPs are narrower than the actual information need of the user. If SERPs are broader then the gradient of the Seen–Selected curve, (g), will be high. Note that α is a case-specific term which affects the maximum gain—it is determined by several factors, such as subjective–specificity of search results and case-specific factors like the search task, and the maximum number of search results the user is expecting to gain.

This model can predict the subjective–specificity of SERPs to the current information need of the user. It allows to compare the gradient of the Seen–Selected graph based on the user's selection behaviour on the current SERP with that of the user's baseline Seen–Selected graph. Such a baseline graph can be constructed by observing the everyday interactions of a user with a search tool. Then, if this user formulates a particular query to explore a topic, the gradient of the new Seen–Selected graph can be compared against the gradient of her baseline graph, and so the system can predict whether the SERPs derived from this query is too narrow or broad for her information-need—and adjust the behaviour of the system accordingly.

3 Empirical Evaluation

In order to empirically validate this model, we conducted a user study where 24 computer science student (masters and PhD level) searched for scientific information in research topics that are not very familiar to them. The task for the participants was to collect scientific articles for a scientific essay writing task in a given topic. We used six experts in six different computer science disciplines to define six unique tasks. The experts defined three search queries in each topic which retrieved SERPs from Google Scholar at three levels of specificity: broad, intermediate, and narrow. Prior to the study, we provided a questionnaire to the participants and made sure that the subjective–specificity of the queries were in comply with the participants' knowledge. We asked the participants to scan these SERPs and click articles that they find useful for the given task. We randomized the order in which they get the tasks. We logged the click interactions and plot the Seen–Selected graph as in Figure 1.

In order to confirm that, in accordance with our model, the gradients of the Seen–Selected curves decrease with the increase of the subjective–specificity of SERPs, and that they follow a natural logarithmic distribution, we analysed the overall distribution of the user information gain over information seen for the three types of SERPs. As our model predicts, the gradient of the Seen–Selected curve decreases as the subjective–specificity SERPs increase (see Table 1).

We used Wilcoxon signed-ranked test to statistically compare the gradients of the predicted models of each type of SERPs. The gradients of the broad SERPs (Mdn 3.56) were significantly greater than the gradients of intermediate (3.08) and narrow SERPS (2.04). The gradients of the predicted models of the intermediate SERPs were significantly greater than that of narrow SERPs.

This empirical evaluation shows that our model captures the effects of subjective–specificity of SERPs to the information-need of the user.

An important future challenge is to investigate in a real exploratory search scenario the performance of the formal model that we developed to predict the subjective–specificity of search results. In the future, we will incorporate our

Table 1. Logarithmic regression models and model fit (R^2) for number of articles Seen–Selected. Breakdown per Broad, Intermediate and Narrow SERPs.

Query	**Model**	**Fit** (R^2)
Broad	$3.83 \ln(n) - 3.59$	0.97
Intermediate	$2.40 \ln(n) - 2.06$	0.97
Narrow	$2.05 \ln(n) - 1.96$	0.97

model in a running IR system and further validate its usefulness in enhancing performance of exploratory search tasks.

References

1. Marchionini, G.: Exploratory search: from finding to understanding. J. Comm. ACM, 49(4), 41–46 (2006)
2. White, R. W., Kules, B., Drucker, S. M.: Supporting exploratory search. J. Comm. ACM, 49(4), 36–39 (2006)
3. Ferragina, P., Gulli, A.:A personalized search engine based on Web-snippet hierarchical clustering. J. Software: Practice and Experience, 38(2), 189–225 (2008)
4. Kelly, D., Fu, X.: Elicitation of term relevance feedback: an investigation of term source and context. In SIGIR, 453–460 (2006).
5. Yee, K. P., Swearingen, K., Li, K., Hearst, M.: Faceted metadata for image search and browsing. In: CHI, (2003)
6. Chau, D. H., Kittur, A., Hong, J. I., Faloutsos, C.: Apolo: making sense of large network data by combining rich user interaction and machine learning. In CHI, pp 167–176 (2011)
7. Kki, M.:Findex: search result categories help users when document ranking fails. In CHI, pp 131–140 (2005)
8. Teevan, J., Alvarado, C., Ackerman, M. S., Karger. D. R.: The perfect search engine is not enough: a study of orienteering behavior in directed search. In CHI, pp 415–422 (2004)
9. Cole, M. J., Gwizdka, J., Liu, C., Belkin, N. J., Zhang, X.: Inferring user knowledge level from eye movement patterns. J. Inf. Proc. Manag (2012)
10. Phan, N., Bailey, P., Wilkinson, R.: Understanding the relationship of information need specificity to search query length. In SIGIR, ACM, pp 709–71 (2007)
11. Joachims, T.: Optimizing search engines using click-through data. In SIGKDD, pp 132–142 (2002)
12. Pirolli, P., Card, S.: Information foraging in information access environments. In: CHI, pp 51–58, (1995)
13. Głowacka, D., Ruotsalo, T., Konyushkova, K., Athukorala, K., Kaski, S., Jacucci, G.: Directing exploratory search: Reinforcement learning from user interactions with keywords. In: IUI (2013)
14. Athukorala, K., Hoggan, E., Lehtio, A., Ruotsalo, T., Jacucci, G.: Information-seeking behaviors of computer scientists: Challenges for electronic literature search tools. In: AGM of Assoc. Inf. Sci. Tech. (2013)