Interactive Visualization of Search Intent for Exploratory Information Retrieval

Antti Kangasräasiö[1] ANTTI.KANGASRAASIO@HIIT.FI
Dorota Głowacka[2] DOROTA.GLOWACKA@HIIT.FI
Tuukka Ruotsalo[3] TUUKKA.RUOTSALO@HIIT.FI
Jaakko Peltonen[1,4] JAAKKO.PELTONEN@HIIT.FI
Manuel J. A. Eugster[1] MANUEL.EUGSTER@HIIT.FI
Ksenia Konyushkova[2] KSENIA.KONYUSHKOVA@HIIT.FI
Kumaripaba Athukorala[2] KUMARIPABA.ATHUKORALA@HIIT.FI
Ilkka Kosunen[2] ILKKA.KOSUNEN@HIIT.FI
Aki Reijonen[2] AKI.REIJONEN@HIIT.FI
Petri Myllymäki[2] PETRI.MYLLYMAKI@HIIT.FI
Giulio Jacucci[2] GIULIO.JACUCCI@HIIT.FI
Samuel Kaski[1,2] SAMUEL.KASKI@HIIT.FI
[1]Helsinki Institute for Information Technology HIIT, Dep. of Information and Computer Science, Aalto University
[2]Helsinki Institute for Information Technology HIIT, Dep. of Computer Science, University of Helsinki
[3]Helsinki Institute for Information Technology HIIT, Aalto University
[4]School of Information Sciences, University of Tampere

Abstract

Searching for relevant documents from a vast amount of scientific data is a challenging problem that requires a close interaction between the user, the search interface and the search engine. This extended abstract summarizes recent research on *Intent Radar*, an interactive search user interface that allows the user to directly interact with her estimated search intent, and in this way direct the search in an intuitive way without the need to type specific queries. In user experiments, Intent Radar improves task performance and quality of retrieved information without compromising the task execution time. This workshop paper presents to the ICML Crowdsourcing and Human Computing workshop audience our work recently published in CIKM 2013 and IUI 2013 as well as a short discussion of ongoing work on the topic.

ICML '14 Workshop: Crowdsourcing and Human Computing, Beijing, China, 2014.

1. Introduction

Exploratory search is a complex task where the user is learning new information while investigating the information space (Marchionini, 2006). For example, researchers face this problem while searching for relevant documents in a field they are not very familiar with.

The exploratory search process is an iterative process and the user's information need evolves based on the search results at each step. A common search strategy in exploratory search is to start with an imprecise query that hopefully leads to the correct part of the information space, and then to direct the search around the initial entry-point in the information space to obtain the information of interest (Teevan et al., 2004).

Current methods for supporting users in exploratory search are either based on suggesting query terms, or allowing faster access to the present search results by faceted browsing or search result clustering (Yee et al., 2003; Hearst & Pedersen, 1996). However, these feedback mechanisms have the disadvantages that they can trap the user to the initial query context and cause cognitive burden to the user (Kelly & Fu, 2006).

Our recent research (Ruotsalo et al., 2013a;b; Głowacka et al., 2013) demonstrates that by allowing the user to directly interact with a model of the estimated search intent, she is able to perform better on information seeking tasks

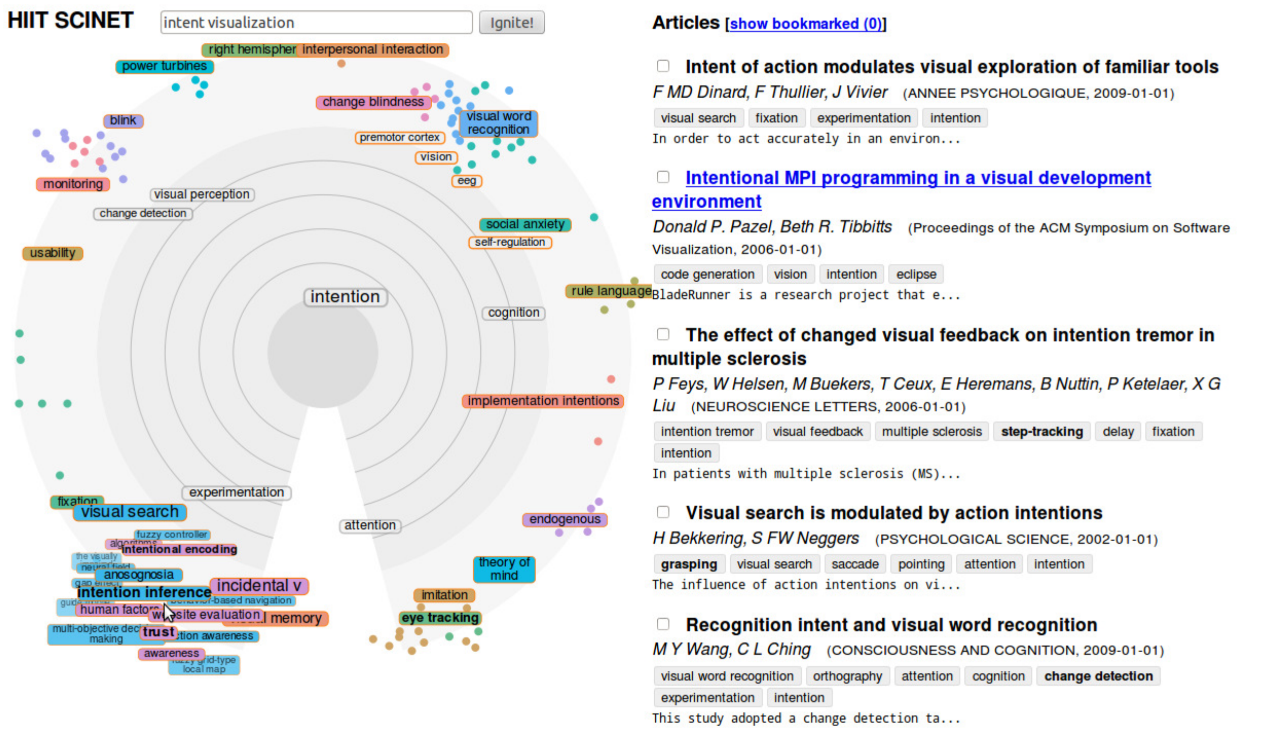

Figure 1. The Intent Radar interface. Search intents are visualized through keywords on a radial layout. The closer a keyword is to the center, the more relevant it is to the estimated interest of the user. Predicted future intents are visualized as keywords in the outer circle. Clustered keywords can be inspected in more detail by using a fisheye lens at the mouse cursor. Relevant documents based on the current search intent are shown as a list on the right.

and find a larger amount of relevant items in a large collection of scientific documents, compared to other baseline systems, such as the popular search interface Google Scholar.

2. Search User Interface

The *Intent Radar* interface is shown in Figure 1. The interface assists users in exploring the information space related to a given information need. The search interface prevents the user from suffering of too much cognitive overload compared to more traditional search interfaces by (1) allowing users to make sense of the information space around the initial query context, and (2) allowing users to intuitive direct the search by interacting with the Intent Radar. Keywords in the inner circle of the Intent Radar represent current interests and keywords in the outer circle represent possible future intents nearby in the information space. The layout is computed by a nonlinear dimensionality reduction algorithm. The algorithm places keywords considered to be more relevant to the user's interests closer to the center of the Intent Radar, while grouping related keywords. More details can be found from Ruotsalo et al. (2013b).

The search session starts by the user entering a textual query to which the search engine returns a set of documents and a set of keywords. The search engine approximates the user intent based on the initial set of the retrieved documents. The user interface presents the user with the intent model (visualized through the Intent Radar) and a list of the most relevant documents. The user then provides relevance feedback for the keywords by dragging a keyword on the Intent Radar to a location representing the desired relevance, i.e. the closer to the centre a given keyword is moved, the more relevant it is. Negative relevance feedback is possible by dragging a keyword outside the radar. After the user lets go of a keyword, the search engine (1) adjusts the intent model based on this feedback and (2) retrieves a new set of documents based on the updated model parameters. This process is iterated until the user is satisfied with the results or decides to start a new search session by typing a new query.

3. Interactive Intent Modeling

Interactive intent modeling relies on two underlying models: the *user intent model* and the *document retrieval*

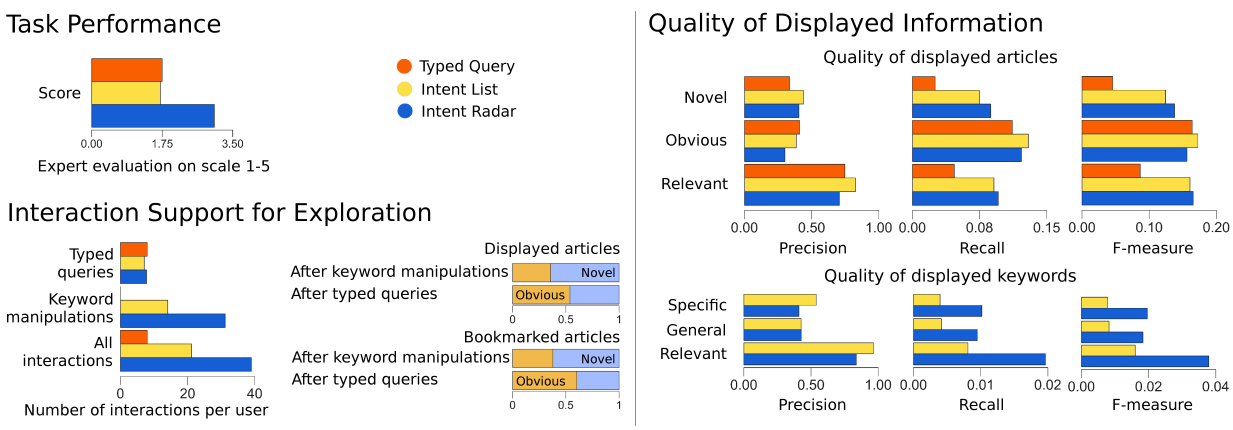

Figure 2. Comparison of the performance of Intent Radar, Intent List (a simplified interface representing intents only as a list of top keywords) and Typed Query (a traditional search interface based on typing queries). **Task performance**: Intent Radar improves users' task performance (answers to predefined questions) compared to other retrieval methods. **Interaction Support for Exploration**: Users interacted with the Intent Radar interface more than with the comparison systems (without compromising the task execution time). The keyword manipulations resulted in the user viewing and bookmarking more novel articles. **Quality of Displayed information**: Intent Radar helps users to move away from the initial query context, thus increasing recall while preserving precision in particular for novel information. Displayed articles: (1) relevance — is this article relevant to the search topic, (2) obviousness — is this a well-known overview article in a given research area, and (3) novelty — is this article an uncommon yet relevant to a given topic or specific subtopic in a given research area. Displayed keywords: (1) relevance — is this keyword relevant for the topic, (2) general — does this keyword describe a relevant subfield, and (3) specific — does this keyword describe a relevant specifier for the subfield?

model. The user intent model estimates the current and future relevance of intent keywords. The document retrieval model is used to search for the most relevant documents based on the current user interests which, in turn, are based on the user intent model.

The search intent is modeled as a multi-armed bandit problem, where each keyword is a "bandit arm" with an expected reward equal to the relevance score provided by the user for that particular keyword. Each keyword is represented as a binary feature vector of length n, where n is the number of documents and each entry in the feature vector indicates presence or absence of a given keyword in a document. The optimal keywords to show to the user (i.e. optimal arms to sample) are selected by the reinforcement learning algorithm LinRel (Auer, 2003). LinRel takes as input the relevance scores given by the user and outputs estimates of the relevance scores for all the keywords, together with error bounds. LinRel orders keywords based on their upper confidence bounds, being roughly the largest relevance score that the keywords would most likely get based on the user's feedback so far and the error bounds. This allows the system to balance between showing the user keywords with large error bounds (exploring the intent model) and keywords with large estimated relevance (exploiting the intent model). Because of this, the user is able to direct her search without being trapped in the initial query context.

The documents are retrieved based on a language modeling approach of information retrieval (Zhai & Lafferty, 2004). The probability of an article to generate the keywords in the user intent model is estimated by a unigram language model with Bayesian Dirichlet Smoothing. Documents are ranked by their probability given the user intent model. To expose the user to more novel documents, Dirichlet Sampling is used to choose the documents presented to the user from the full ranked list.

4. Main Results

The ability of the Intent Radar to enable better task performance has been studied in task-based experiments with 30 university students. The task was to (1) bookmark scientific articles relevant to a certain topic, and (2) answer a set of predefined questions related to the topic. The users had to search for this information from a set of 50 million scientific articles using one of three possible interfaces: the Intent Radar (see Figure 1); Intent List, which is similar to Intent Radar, but with the radar layout replaced with a list of the most relevant keywords; and Typed Query, where the user has to type in textual queries at each iteration. More detailed description of the experimental design can be found in Ruotsalo et al. (2013b); Głowacka et al. (2013). Another study, where Intent Radar was compared against Google Scholar, is presented in Ruotsalo et al. (2013a).

4.1. Summary of Main Results

A summary of the experimental results from Ruotsalo et al. (2013b) is presented in Figure 2. The users were randomly assigned one of three interfaces and asked to answer a set of questions from a given topic. All the participants completed the task within the given timeframe of 30 minutes. There were no significant time differences between the systems or tasks. The user's task performance was measured by the grades given by experts to the answers to the predefined questions (1 being the worst and 5 being the best). The system with the Intent Radar interface noticeably improved user's task performance (Figure 2, Task Performance). The Intent Radar interface also enhanced users' interaction with the system, as users engaged in exploratory search for up to three times more iterations compared to simple query typing (Figure 2, Interaction Support for Exploration, left). Additionally, the interfaces that allowed keyword manipulation presented the user with more novel documents, which supports the exploration of the information space (Figure 2, Interaction Support for Exploration, right). Most importantly, interactions with the Intent Radar resulted in improved quality of retrieved information (Figure 2, Quality of Displayed Information). The recall and F-measure of novel and relevant documents returned by the interfaces supporting interaction with the user model were noticeably higher than with the traditional query interface. Also, the keywords displayed by the Intent Radar had noticeably higher recall and F-measure in all keyword classes compared to the Intent List.

4.2. Future Directions of Research

Although the experimental results clearly show that the Intent Radar provides a better support for the user in exploratory type search tasks compared to more traditional interfaces, post-experimental interviews with the participants (Głowacka et al., 2013) indicate that there is still room for improvement. The majority of the feedback was positive, however, some of the users indicated that, for example, the changes in the keyword display in the Intent Radar were too rapid and sometimes unexpected, or that there was no option to go back to a previous iteration and change the given feedback.

Based on this user feedback, we are currently exploring new ways of visualising keywords on the Intent Radar that will allow the user to have more control over the search process and to have a better understanding of the consequences of her actions, i.e. keyword movements, on future search results. Possible improvements include applying various approximation techniques to allow the user to see what the next search iteration might look like given a particular keyword movement, or making the transition from one iteration to another smoother and more gradual. An additional line of research is using implicit feedback from physiological signals, such as EEG, heart rate and skin conductance in the search.

Acknowledgments

This work has been partly supported by the Academy of Finland (255725, Finnish Centre of Excellence in Computational Inference Research COIN, 251170, 252845, 256233) and TEKES (D2I). Certain data included herein are derived from the Web of Science prepared by THOMSON REUTERS, Inc., Philadelphia, Pennsylvania, USA: Copyright THOMSON REUTERS, 2011. All rights reserved. Data is also included from the Digital Libraries of the ACM, IEEE, and Springer.

References

Auer, P. Using confidence bounds for exploitation-exploration trade-offs. *The Journal of Machine Learning Research*, 3:397–422, 2003.

Głowacka, D., Ruotsalo, T., Konyushkova, K., Athukorala, K., Kaski, S., and Jacucci, G. Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Proceedings of IUI 2013, International Conference on Intelligent User Interfaces*, pp. 117–128, New York, NY, 2013. ACM.

Hearst, M. A. and Pedersen, J. O. Reexamining the cluster hypothesis: Scatter/gather on retrieval results. In *Proceedings of the 19th Annual ACM International Conference on Research and Development in Information Retrieval (SIGIR 1996)*, pp. 76–84, New York, NY, 1996. ACM.

Kelly, D. and Fu, X. Elicitation of term relevance feedback: An investigation of term source and context. In *Proceedings of the 29th Annual ACM International Conference on Research and Development in Information Retrieval (SIGIR 2006)*, pp. 453–460, New York, NY, 2006. ACM.

Marchionini, G. Exploratory search: From finding to understanding. *Communications of the ACM*, 49(4):41–46, 2006.

Ruotsalo, T., Athukorala, K., Głowacka, D., Konyushkova, K., Oulasvirta, A., Kaipiainen, S., Kaski, S., and Jacucci, G. Supporting exploratory search tasks with interactive user modeling. In *Proceedings of ASIST 2013, the 76th ASIS&T Annual Meeting*, Silver Spring, MD, 2013a. Association for Information Science and Technology.

Ruotsalo, T., Peltonen, J., Eugster, M. J. A., Głowacka, D., Konyushkova, K., Athukorala, K., Kosunen, I., Reijonen, A., Myllymäki, P., Jacucci, G., and Kaski, S. Directing exploratory search with interactive intent modeling. In *Proceedings of CIKM 2013, the ACM International Conference of Information and Knowledge Management*, pp. 1759–1764, New York, NY, 2013b. ACM.

Teevan, J., Alvarado, C., Ackerman, M. S., and Karger, D. R. The perfect search engine is not enough: A study of orienteering behavior in directed search. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2004)*, pp. 415–422, New York, NY, 2004. ACM.

Yee, K. P., Swearingen, K., Li, K., and Hearst, M. Faceted metadata for image search and browsing. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2003)*, pp. 401–408, New York, NY, 2003. ACM.

Zhai, C. and Lafferty, J. A study of smoothing methods for language models applied to information retrieval. *ACM Transactions on Information Systems*, 22(2):179–214, 2004.