Towards Inferring Web Searcher Intent from Behavior Data

Eugene Agichtein
Emory University
Mathematics & Computer Science Department
Atlanta, GA 30322
eugene@mathcs.emory.edu

Qi Guo
Emory University
Mathematics & Computer Science Department
Atlanta, GA 30322
qguo3@emory.edu

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract

Inferring searcher intent is a central problem in information retrieval: for effective ranking and result presentation, the retrieval system must know what the searcher is looking for. This position paper describes our ongoing efforts towards the development of machine learning techniques to infer search *intent* from search behavior *actions* such as query reformulations, clicks, browsing, or scrolling over the search results. That is, our goal is to develop the methodology to combine two sides of the search process, query specification and result interactions, by training robust machine learning models to recover the search intent from observable behavior clues.

Workshop Areas: Methods; Models

Introduction

Web search and related activities have transformed society and now are the primary methods of accessing and discovering information. Inferring searcher intent, and translating it into effective retrieval, is perhaps the key fundamental problem in information retrieval that is still not solved. Sadly, current methods for inferring search intent are inadequate. The user is expected to formulate the initial query (possibly without knowledge of the underlying collection vocabulary or domain terminology), and then repeatedly re-formulate the query until the needed documents are retrieved. Deployed IR systems largely ignore the *context* of the search with a search session, as well as the *behavior* of searchers interacting with the results.

The goal of our work is to change this situation. Imagine a search engine that could "read your mind": based on search context, previous interactions of other users, and the current user's searching behavior, the system would infer the *intent* behind the search. After retrieving the initial set of results, the search engine would automatically detect whether the results returned are satisfactory, and adjust the ranking and presentation accordingly, or perhaps suggest queries that would uncover other aspects of the inferred intent. For this, the search engine would analyze the parts of the documents viewed, and any explicit or implicit indications of interest (e.g., highlighting parts of document, mouse movement and position, or scrolling).

To achieve this vision, we are investigating robust and scalable data mining and machine learning techniques for analyzing observable searcher behavior and interaction data, which can provide crucial clues about the searcher intent and state of mind. For example, rent work demonstrated some coordination between gaze position (and therefore, interest) and mouse position [3]. The rest of this paper describes our models, and reports some of our recent results.

Search Model

Our work starts with a hierarchical model of search, following recent information behavior literature. We assume that a user is attempting to accomplish a specific (directed) search task by solving specific search goals, as illustrated in Figure 1. Many tasks require

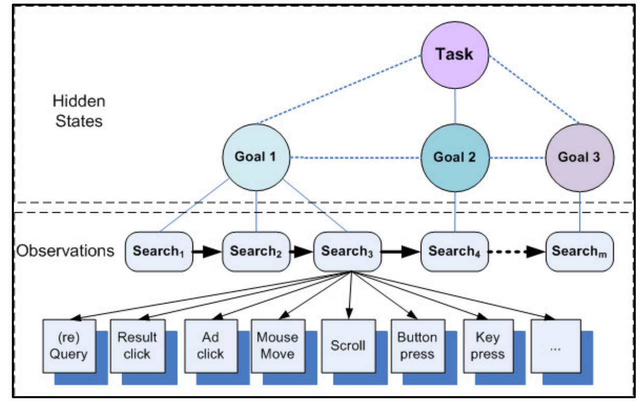


Figure 1: Search model

multiple searches until the needed information is found. For each search, the user interacts with the Search Engine Results Pages (SERPs) and the "landing" pages (the click destination documents). The interactions such as result clicks, mouse movements, scrolls, or key presses, together with the corresponding SERP or landing page content, form the observations that can be captured directly with proper instrumentation of the browser and/or the SERP pages.

Capturing Interaction Data

In order to collect realistic searching behavior data, we have developed an instrumented version of the LibX toolbar for the Firefox browser. The instrumented browsers are deployed on roughly 150 public-use machines in the Emory University Libraries, where our instrumentation and data capture has been running since January 2008. The search behavior of users (by explicit opt-in mechanism) is tracked by saving mouse movements, clicks, page content, and all necessary information to re-construct the searcher's behavior. This instrumentation has been approved by the Emory IRB, and so far more than 5,000 users opted in, with roughly 1,000 search sessions via Google, Google Scholar, and the Emory search engines, EUCLID and Discover-E logged daily.

Machine Learning Framework

Our goal is to predict intent (e.g., each goal) in a search session, given a sequence of *observations*. That is, the search intent at step *j* in the search session will be predicted given a sequence of observations of previous and current actions $CA_1, \ldots, CA_j$. This learning task can be naturally modeled as *sequential labeling*, using graphical models such as Conditional Random Fields (CRFs), illustrated in Figure 2.

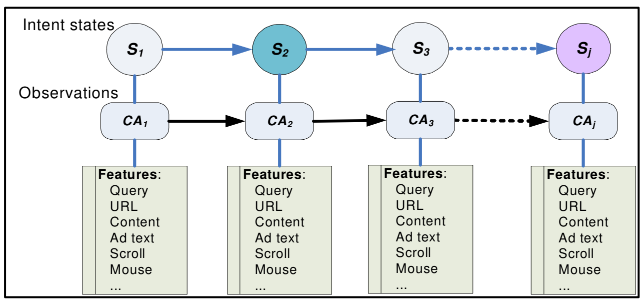


Figure 2: An example Conditional Random Field (CRF) configuration for intent inference

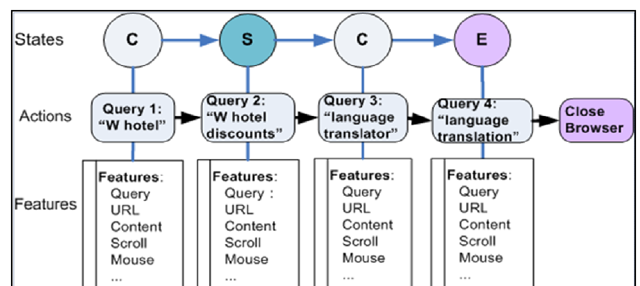
The sequence of context and action observations $\{CA\}$ are *features*, and the sequence of search intents $\{S\}$ are the *hidden states* to predict. With this formulation, the specific problems and search tasks we address can be reduced to answering four basic questions: (1) what are the possible target values of *topic* and *mental state* classes; (2) how to represent the context and action observations; (3) how to characterize and learn the correspondence between observations and hidden states, and (4) how to realistically evaluate the resulting models. We now describe example applications to illustrate our methodology.

Example Applications

We consider two search tasks, both directly connected to intent detection: predicting switch in the searcher's goal, and inferring user interest in search advertizing.

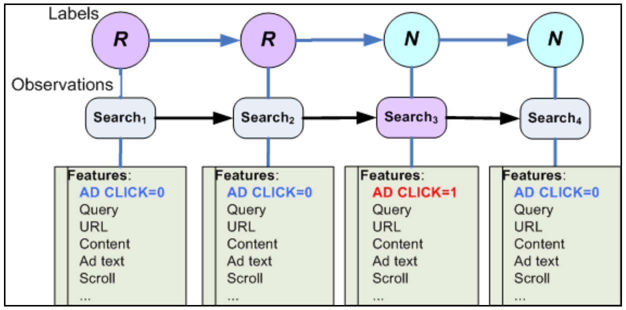
Predicting switching in search goal

Effective search session segmentation -- grouping queries according to common task or intent -- can be useful for improving relevance, search evaluation, and query suggestion. Our framework allows for predicting, in real time, whether a user is about to switch interest -that is, whether the user is about to finish the current search and switch to another search task (or stop searching altogether). The corresponding CRF model is shown below. The hidden states are *switch goal* (S), *continue with current goal* (C), and *end search* (E). The task, then, is to predict the *next* state in the search, given the previous states and actions in the search session. As reported in reference [1], the overall state prediction performance (macro-averaged F1) using our rich behavior model is 0.49, compared to 0.25 when using query text information alone.



States (C-Continue; S-Switch; E-Exit), actions, and features model for an example user session consisting of two search missions

Inferring user interest in search advertizing
The goal is to predict the current searcher's interest in commercial advertising for the query – that is, whether the searcher is currently "receptive" to advertising. The states in the model (shown below) are *receptive* (*R*) and *non-receptive* (*N*).



In this model, if a searcher is in state *R*, she is predicted to click on an ad result sometime *in the future* within the same search session; if a searcher is in a state *N*, she is predicted *not* to click on an ad result in the same session. These predictions could be useful to decide how many and which ads to show, or not to show ads at all for *this* searcher (e.g., to avoid annoying her if she is not in a receptive state). As described in reference [2], our CRF-based method was able to improve *future* ad clickthrough prediction precision by over 200% relative to a commercial search engine and current state-of-the-art techniques, at about 22% coverage (recall) level where sufficient interaction evidence was available.

Design Implications

"One size fits all" is well known as a drawback for one universal design of a system. We believe that user interface designs (specifically, search interfaces) could be further improved with our methods.

First, *session-level search experience* could be improved with a more intelligent user interface that dynamically adjusts the search result ranking and presentation based on the searcher's implicit feedback and interactions. Examples of possible functionality of such an interface include real-time, contextual re-ranking of the search results to improve relevance; or intervening when users appear to be get frustrated with the search results (e.g., by more aggressive query suggestion). Second, *search interface design* could be improved by developing more fine-grained, effort-based evaluation metrics for evaluating search engine result quality, and result presentation effectiveness. That is, our models could be trained to detect patterns of frustrated behavior associated with poor result presentation, distinct from, say, frustration caused by poor search result quality.

Beyond the web search setting, our techniques could be also applicable to recommendation systems, and online social networking tools - or other information-centric systems that rely on extensive user interactions.

References

[1] Guo, Q., and Agichtein, E: Beyond session segmentation: predicting changes in search intent with client-side user interactions. *In SIGIR 2009*: 636-637

[2] Guo, Q., Agichtein, E., Clarke, C., and Ashkan, A. In the Mood to Click? Towards Inferring Receptiveness to Search Advertising. *In Web Intelligence 2009*: 319-324

[3] Rodden, K. , Fu, X., Aula, A., and Spiro, I. Eye-mouse coordination patterns on web search results pages, In *Ext. Abstracts, CHI 2008*