



Understanding and Modeling Users of Modern Search Engines

A. Chuklin

Understanding and Modeling Users of Modern Search Engines

As search is being used by billions of people, modern search engines are becoming more and more complex. And complexity does not just come from the algorithms. Richer and richer content is being added to search engine result pages: news and sports results, definitions and translations, images and videos. Many such elements are added by search engines in their attempt to stand out from the competition by providing a superior user experience. However, the more complex search engines become, the harder it gets to understand users and their interactions with result pages, and to measure the quality of the user experience. In this thesis we address exactly this topic.

We start by analyzing user behavior on complex result pages and show that the users' click patterns are non-trivial. We also demonstrate that there are situations where we observe no clicks at all, even though there is good content on a search engine result page and the users are likely to be satisfied. Having made these observations, we proceed to *modeling* the users of modern search engines. In particular, we suggest so-called *intent-aware* models that capture different user intents and different presentation types on the search engine result page. We also contribute to the area of user model evaluation by presenting a comprehensive evaluation of click models and a new evaluation method designed specifically for modern search engine result pages.

We then turn our attention to *evaluating* the user search experience. With additional tools made available by modern search engines, users hope to become more productive in their everyday life and achieve more in a smaller amount of time. That can reliably be achieved only if the added complexity of the search engines comes with a correctly measured benefit. We suggest a general framework for deriving an offline evaluation metric from a click model—*click model-based metrics*—and then refine the underlying user model by modeling user attention and user satisfaction signals in addition to clicks. The benefit of such evaluation metrics is that they can be fit to real user data and then re-used multiple times without involving actual users.

Next to studying offline metrics, we also study online evaluation approaches that require deploying new versions of the search system live and comparing them to the current production system based on user feedback. We propose a modification of a popular online evaluation method—interleaving—that accounts for blocks of rich content (often called verticals) that we refer to as *vertical-aware interleaving*. In addition, we suggest several methods for evaluating evaluation methods, both offline and online ones, from multiple angles.