

# Estimating the Query Difficulty for Information Retrieval

David Carmel, Elad Yom-Tov

IBM Research lab in Haifa,  
{carmel,yomtov}@il.ibm.com

## Abstract:

Many information retrieval (IR) systems suffer from a radical variance in performance when responding to users' queries. Even for systems that succeed very well on average, the quality of results returned for some of the queries is poor. Thus, it is desirable that IR systems will be able to identify "difficult" queries in order to handle them properly. Understanding why some queries are inherently more difficult than others is essential for IR, and a good answer to this important question will help search engines to reduce the variance in performance, hence better servicing their customer needs.

The high variability in query performance has driven a new research direction in the IR field on estimating the expected quality of the search results, i.e. the query difficulty, when no relevance feedback is given. Estimating the query difficulty is a significant challenge due to the numerous factors that impact retrieval performance. Many prediction methods have been proposed recently. However, as many researchers observed, the prediction quality of state-of-the-art predictors is still too low to be widely used by IR applications. The low prediction quality is due to the complexity of the task, which involves factors such as query ambiguity, missing content, and vocabulary mismatch.

The goal of this tutorial is to expose participants to the current research on query performance prediction (also known as query difficulty estimation). Participants will become familiar with states-of-the-art performance prediction methods, and with common evaluation methodologies for prediction quality. We will discuss the reasons that cause search engines to fail for some of the queries, and provide an overview of several approaches for estimating query difficulty. We then describe common methodologies for evaluating the prediction quality of those estimators, and some experiments conducted recently with their prediction quality, as measured over several TREC benchmarks. We will cover a few potential applications that can utilize query difficulty estimators by handling each query individually and selectively based on its estimated difficulty. Finally we will summarize with a discussion on open issues and challenges in the field.

## ACM Categories & Descriptors:

H.3.3 [Information Search  
and Retrieval]: Retrieval models

## General Terms:

Algorithms, Measurement

## Keywords:

Retrieval robustness, Query difficulty estimation, Performance prediction

## Bio/Bios:

**David Carmel** is a Research Staff Member at the Information Retrieval group at IBM Haifa Research Laboratory. David earned his PhD in Computer Science from the Technion, Israel Institute of Technology in 1997. David's research is focused on search in the enterprise, query performance prediction, social search, and text mining. For several years David taught the Introduction to IR course at the CS department at Haifa University.

At IBM, David is a key contributor to IBM enterprise search offerings. David is a co-founder of the Juru search engine which provides integrated search capabilities to several IBM products, and was used as a search platform for several studies in the TREC conferences. David has published more than 60 papers in Information retrieval and Web journals and conferences, and serves in the Program Committee of many conferences (SIGIR, WWW, WSDM, CIKM, ECIR) and workshops.

**Elad Yom-Tov** is a Research Staff Member at the Analytics Department at the IBM Haifa Research Laboratory. The main focus of his work is research into methods for large-scale machine learning, with a recent focus on social analytics. Prior to his current position he worked at Rafael Inc., where he applied machine learning to image processing. Elad is a graduate of Tel-Aviv University (B.Sc.) and the Technion, Haifa (M.Sc. and Ph.D). He is the author (with David Stork) of the Computer Manual to accompany Pattern classification, a book and a Matlab toolbox on pattern classification.

Elad's work in Information Retrieval includes query difficulty estimation, social tagging, and novelty detection.

**Both** David and Elad published many papers on query performance prediction, and organized a workshop on this subject in SIGIR 2005. Their paper on learning to estimate query difficulty won the Best Paper Award at SIGIR 2005.