The World-Wide Quest for Better Web Search

The World Wide Web is an ever-growing sea of information. The volume of data available online increases exponentially, as more and more people and entities publish data on the Web. In this chaotic arena of text, images, voice and video, hundreds of millions of people from all over the world are surfing in search of information on every possible walk of life.

Web search is made possible by complex information retrieval systems called search engines, which combine (among others) such disciplines as database technology, distributed computing and storage, statistical linguistics and graph algorithms. Users submit queries to the engines, and expect their queries to be instantaneously answered with ranked lists of the most relevant URLs available online for each query. In order to meet these demands, search engines must collect and index billions of resources, and develop highly efficient retrieval and ranking algorithms that are capable of effectively answering queries in almost a blink of the eye.

This talk will highlight several challenges in search engine technology and Web structure research, such as applications of link analysis, the war against spam, e-commerce recommender systems and power-law phenomena on the Web. Time permitting, some aspects of the business side of Web searching will also be covered.

The talk is self contained and does not assume any special background.