Yale University Computer Science Talk Sublinear Algorithms for Massive Datasets

Host: Yang Cai



Erik Waingarten

Wednesday - February 16, 2022 4:00 p.m.

Zoom Presentation

Abstract:

The influx of massive data systems poses a unique challenge for algorithms research. In this modern era, the classical theory of algorithms is often insufficient, and our goal is to develop the theory of sublinear computation.

This talk will cover various scenarios where the resources used by an algorithm (running time, memory, number of measurements, ... etc) should be significantly smaller than the input size. We will spend most of the time on *sublinear time* algorithms for similarity search in high-dimensional spaces and *sublinear space* algorithms for the optimal transport problem, where we will present new algorithmic and analytical techniques for tackling these questions. A central theme throughout the talk is the notion of randomized space partitions, and how they lead to algorithms which are simple, have provable guarantees, and are extremely useful.

Bio:

Erik Waingarten is an NSF Postdoctoral Fellow at Stanford University's Computer Science department hosted by Moses Charikar. Prior to that, he was a research fellow at the Simons Institute for the Theory of Computing. He obtained his PhD from Columbia University, where he was advised by Xi Chen and Rocco Servedio. He is interested in the design and analysis of algorithms for massive datasets, with a particular focus on the interplay of high-dimensional geometry, sketching, and property testing.

