CS 590 Sublinear algorithms Jan 11, 2021

Syllabus

Lecturer: Elena Grigorescu

1 Schedule

Tu/Th 1:30-2:45 pm (on Zoom. See link on Brightspace)

2 Instructor

Elena Grigorescu, elena-g@purdue.edu, office hours by appointment.

3 Tech Platforms

We'll be using Brightspace to post Zoom lecture recordings and homework. Check there the invite for Campuswire, which is our discussions platform.

4 Description

Algorithms that compute with a sublinear amount of resources (time, space, communication) lie at the foundations of data science. They can be used to analyze data across domains that span from social networks, financial transactions, or genomics data, to satellite communications or astronomical data. In this course we will study several sublinear models, including the sublinear-time model of Property Testing, sublinear-space streaming models, and sublinear-communication models. We will introduce many of the combinatorial, algebraic and geometric techniques commonly used in analyzing very fast algorithms for mathematical objects such as graphs, strings, distributions, and error-correcting codes.

Prerequisites: Discrete math and mathematical maturity.

5 A partial list of topics

Sublinear time approximation algorithms for average degree, number of connected components, MST, clustering. Property testing: Testing triangle freeness in dense graphs (Szemeredi's regularity lemma, The "triangle removal lemma".) Yao's principle for showing lower bounds. Streaming algorithms for graph problems -e.g., connectivity; bipartiteness; distance; matchings Locally testable/decodable codes (Hadamard, Reed-Muller, matching vector codes). Communication complexity -eg, connectivity, min-cut, sparsification; lower bounds for streaming and property testing. Distributed learning theory.

6 The project

The project will be either a research project or an in-depth survey of a topic of your choice. You will need to submit a topic, several partial progress reports, and a final report plus presentation. I will be guiding you through all these stages. The project can be done in groups of at most 2 people. To get started, pick a relatively recent paper on sublinear algorithms from a theoretical CS conference (FOCS, STOC, SODA, CCC, APPROX-RANDOM, ICALP, ICS). Some good places to start from are the Electronic colloquium on computational complexity and the Algorithms and data structures arxiv.

Timeline:

Jan 19: start brainstorming for project ideas, possibly with one collaborator.
by Feb 2: run your ideas by me and I'll help you choose a topic.
Feb 16: submit 1-2 paragraph proposal
March 23: 5-10 mins partial progress presentation to class
April 29: project outcomes presentation
April 30: report submission

7 Grading

4-5 homework sets 40%
Semester long research project 45%
Scribe notes for 2-3 lectures + peer homework grading 10%
Class participation 5%