Graphs and Networks Lecture 1

Course Description and Mechanics

Daniel A. Spielman September 2, 2010

1.1 Registration

If you plan to take this class, please sign up on the Classes V2 server. This is how you will receive class-related announcements.

Conversely, if you do not plan to take this class, please remove yourself on the server.

1.2 Overview

The purpose of this class is to introduce many types of graphs that are studied in the sciences and engineering, to examine some elementary processes that take place on those graphs, and to introduce standard analyses of graphs. "Network" is just another word for "graph".

This course will combine mathematical and experimental approaches. Some of the lectures will consist of the presentation of definitions, theorems and proofs. Others will present observations and experiments.

The course will be divided into three sections. In the first section, we will learn what graphs look like. We will examine graphs from many different sources, including social networks (who knows who), the web graph (which page links to which), the internet graph (which router links to which), citation graphs (who references whose papers), protein interaction networks (which proteins interact with each other), planar graphs (which country is next to which), well-shaped meshes (pretty pictures with triangles), geometric graphs (who is near who), random graphs (whichever), random power-law graphs, and algebraically defined graphs (links determined by mathematician). We will make fundamental measurements of the real-world graphs, and will use these to compare them to abstract models of graphs. We will see that all models are bad, but that some are worse than others. We will also see that some measurements of real-world graphs are actually artifacts of the way these graphs were obtained.

In many applications, people are interested in processes that occur on the vertices or edges of a graph. Others consider processes on graphs as a means of understanding the structure of a graph. In the second section of the course, we will study many nice properties on graphs, as well as a few counter-intuitive ones. We will focus on random walks (a toddler exploring a graph), diffusion (the toddler makes a mess), contagion and coordination (spread of diseases on graphs), electrical flow (treating edges as resistors), gossip (spreading rumors), traffic flow (and the price of anarchy), flocking (coordination on changing graphs), and the evolution of graphs.

In the final section of the course, we will examine standard analyses of graphs. The most famous is PageRank, a means of evaluating the importance of vertices in a graph. We will also examine measures of the similarity and distance between vertices in graphs, the problem of detecting communities or clusters of vertices, hierarchical decompositions of graphs, graph drawing, the problem of predicting which edges may appear later in a graph, the problem of detecting incorrect edges or non-edges in noisy graph data, and learning and regression on graphs. We will take three approaches to justifying these analyses: telling a good story, proving that they satisfy desirable axioms, and demonstrating that they predict something meaningful.

1.3 Required background

To take this class, one should probably have had courses in discrete mathematics, probability, and linear algebra. Some more advanced courses in mathematics, statistics, or computer science would help. I will try to cover background material as it is needed.

1.4 Course Materials

I have not ordered a book for this class, since I cannot find a good book that covers even a small fraction of the material we will discuss. Instead, I will place links to readings and my lecture notes on the course homepage: `http://www.cs.yale.edu/homes/spielman/462/`.

1.5 Work for the course

There will be 5 or 6 problem sets during the semester. Each problem set will consist of two parts, a theoretical part and an experimental part. Students have the option of doing either just the theoretical part, or a portion of the theoretical part and all of the experimental part. I think of the course as having two tracks: theoretical and experimental.

The students pursuing the experimental track will first have to get hold of a big graph. In fact, it should be more than a graph. There should be some information available about the meaning of every vertex or edge. Students will then use these graphs for experiments throughout the course.

Students have the option of doing course projects instead of some of the problem sets. Typical projects will be worth 1 or 2 problem sets, depending on the difficulty and the number of students collaborating. The details are to be negotiated with me.

1.6 Motivation

I believe that Graph Theory is the new Calculus. Researchers in many fields of science, engineering, and social science have spent much of the last decade actively applying graph theory. My goal in this class is to introduce you to some of these applications, and to give you the background you will need to understand the rest. There are more applications than I can possibly cover, so I'm just going to pick those I understand best.

Networks are very trendy these days, and accordingly inspire bad research as well as good. They provide very useful analogies and analyses, but have also been used to justify many bad ideas and hypothesis. This course should help students figure out which research is which.

This class is designed in opposition to standard classes in graph theory, which are typically based on a book originally written in the 1970's, and which makes no reference to uses of graph theory outside of combinatorics and theoretical computer science.

1.7 Warning

I use the words "vertex" and "node" interchangeably.