Seminar on Combinatorial Computing
February 15, Wednesday, 6:30 p.m.
Room 6417, Graduate Center
365 Fifth Avenue, New York

Ranking and clustering: aggregating inconsistent information

Nir Y. Ailon
Princeton University

Abstract

A ranking of n web pages is to be chosen from outputs of k search engines. How do we choose one ranking minimizing the "disagreement" with the k rankings? A clustering of n genes is to be chosen from outputs of k clustering algorithms. How do we choose one clustering minimizing the "disagreement" with the k clusterings? These information aggregation problems date back to as early as 1785, when the French philosopher Condorcet considered voting systems where each voter chooses a full ranking of a set of candidates. Recently, their computational aspects have been studied.

In this talk, I will describe new algorithms improving on the best known approximation ratios for both the ranking and the clustering problems with respect to a standard measure of

disagreement. I will also discuss related graph theoretic optimization problems, known as the minimum feedback arc set in tournaments and the correlation clustering in complete graphs. Time permitting, I will also talk about a recent result with Moses Charikar about a generalization of Correlation Clustering with applications to phylogeny and hierarchical classification of data, and some interesting related open problems.

This is joint work with Moses Charikar and Alantha Newman.

For further information contact Janos Pach at pach@cims.nyu.edu,
or visit our website
http://www.math.nyu.edu/p̃ach/public_html/combinatorics_seminar.html