A general approach for incremental approximation and hierarchical clustering

David P. Williamson

`dpw@cs.cornell.edu`

School of Operations Research & Industrial Engr.
Cornell University
236 *Rhodes Hall*
Ithaca, NY 14853
USA

Abstract

We consider the incremental k-median problem introduced by Mettu and Plaxton and the hierarchical median problem introduced by Plaxton which are closely related to the well-known k-median problem. Given a metric space with facilities and clients, the objective of the k-median problem is to select a set of facilities to open so as to minimize the sum of distances from the clients to the nearest open facility. In the incremental k-median problem, we require an ordering of facilities so that the cost of opening the first k facilities in the ordering is close to the optimal solution with k facilities for all k. The hierarchical median problem asks for solutions for all k satisfying some "hierarchical" properties such that the cost the solution with k facilities is close to the optimal k-median solution with k facilities. The maximum ratio between the cost of the solution with k facilities and that of the optimum k-median solution over all k is called the competitive ratio.

We give a deterministic 16-competitive and randomized $4e$-competitive algorithm for the incremental k-median problem improving the previously known 29.86-competitive algorithm by Mettu and Plaxton. We also give a deterministic 41.42-competitive and randomized 20.06-competitive algorithm for the hierarchical median problem improving the previously known 238.8-competitive algorithm by Plaxton. Our algorithms are based on sorting approximate solutions for the k-median problem into buckets of geometrically increasing values and "nesting" specific solutions to obtain the incremental solution. Our algorithmic ideas can be generalized to give incremental algorithms for many problems like k-MST, k-vertex cover, uncapacitated facility location problems which satisfy certain basic properties.

This is joint work with Guolong Lin, Chandrashekar Nagarajan, and Rajmohan Rajaraman.