

Exercises, Algorithm Design II, Fall 2012

Rasmus Pagh

5. november 2012

The case studies consider data sets that can be found via the course home page. The class will split into smaller groups that work on one of the cases, and later each group should present their algorithm design. Implementing the solution after the exercises is encouraged, and could yield components useful for the final project. We again consider the point sets in <http://www.tsp.gatech.edu/world/countries.html>.

Algorithm Design Case Study: Two-dimensional median

Data structures such as quad-trees are used to index point collections. A subtask when constructing a quad-tree is to compute a “2D median” point (x, y) that is “median” in the sense that it splits the point set as evenly as possible into four subsets, according to whether they have coordinate values smaller than or larger than x resp. y . Assume for simplicity that two points never have the same x or y value.

1. Give a precise formulation of the problem to be solved. (If you see several equally good formulations, this is also fine.)
2. Suggest a randomized algorithm for computing (your formulation of) the 2D median.
3. Analyze the expected running time of your algorithm and compare to direct deterministic approaches.

Algorithm Design Case Study: Taxicab ripoff

The task is to construct a long (but not necessarily longest) tour that passes n towns. One approach to this is a *random walk*, i.e., randomly choose the next town on the tour among the towns not already visited.

1. Analyze the expected length of the tour found. The answer will be a function of the pairwise distances among cities. (**Hint:** Linearity of expectation.)
2. Can you implement the algorithm such that it runs in $O(n)$ time?
3. Implement the algorithm and run it 10 times on a large point set. How do the lengths of the tours found compare to your analysis (and to the optimal TSP tour)?