

Algorithms on Evolving Data Sets

Mohammad Mahdian, Google INC.

In many real-world applications, an algorithmic problem needs to be solved on an input data set that is constantly changing. For example, search engines need to find a ranking of web-pages based on a changing web-graph. We formulate and study a new stochastic model for computation on evolving data sets. In this framework the input data changes gradually over time and the algorithm can only track these changes by explicitly probing the data set. The goal is to find a compute an output that is close to the correct output with a limit on the number of probes in each time step. We apply this framework to the problems of sorting and selection as well as several graph problems such as connectivity, spanning trees, and PageRank computation, and obtain close-to-optimal algorithms in each case.