

Introduction to data streams

Sukhendu Chakraborty
Department of Computer Science
Duke University

February 14, 2006

Computing over data streams is a recent phenomenon that is of growing interest in many areas of computer science, including databases, computer networks and theory of algorithms. In this scenario, it is assumed that the algorithm sees the elements of the input one-by-one in arbitrary order, and needs to compute a certain function of the input. However, it does not have enough memory to store the whole input. Therefore, it must maintain a sketch of the data. Designing a sketching method for a given problem is a novel and exciting challenge for algorithm design.

In this talk, I will give a high level picture of currently popular data stream models, motivating scenarios and other applications for algorithms in these models. I would try to abstract mathematical ideas, algorithmic techniques and approaches for data stream models which comprise the foundation of the theory of data streams that is emerging.