Research Description

Designing data structures and algorithms with applications to data streaming, energy-efficient computing and social network analysis

The objective is to devise algorithmic techniques to deal with massive data sets that cannot be stored in the random-access memory (RAM) of a computer. This is because the typical size of the RAM of a computer is in the order of Gigabytes. Truly massive amount of data (in the range of Terra-bytes to Peta-bytes) can only be stored in external memory devices but that they are slow, e.g., accessing the data stored in a location requires a few milliseconds. On the other hand, the locations of a RAM can be accessed in a few nanoseconds only. Therefore, the main bottleneck in designing massive data set algorithms lies in the number of I/Os required between the RAM and the external memory devices. The purpose of this project is to design I/O-efficient algorithms, i.e., algorithms that reduces the number of I/Os between the RAM and external memory. In addition, we are also interested to design energy-efficient algorithms to process massive data sets.

Selected Publications


