DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF SOUTHERN DENMARK, ODENSE

COMPUTER SCIENCE COLLOQUIUM

The Input/Output Complexity of Triangle Enumeration

Francesco Silvestri

Università di Padova and ITU Copenhagen

Tuesday, 21 January, 2014 at 14:15

Auditorium U49C

Abstract:

In this talk, we consider the well-known problem of enumerating all triangles of an undirected graph. Our focus is on determining the input/output (I/O) complexity of this problem. Let E be the number of edges, $M < E$ the size of internal memory, and B the block size. The best previous results are $\text{sort}(E^{3/2})$ I/Os (Dementiev, PhD thesis, 2006) and $O(E^2/(MB))$ I/Os (Hu et al., SIGMOD 2013), where $\text{sort}(n)$ denotes the number of I/Os for sorting n items. We improve the I/O complexity to $O(E^{3/2}/(\sqrt{M}B))$ expected I/Os, which improves the previous bounds by a factor $\min(\sqrt{E/M}, \sqrt{M})$. Our algorithm is cache-oblivious and also I/O optimal: We show that any algorithm enumerating t distinct triangles must *always* use $\Omega(t/(\sqrt{M}B))$ I/Os, and there are graphs for which $t = \Omega(E^{3/2})$. Finally, we give a deterministic cache-aware algorithm using $O(E^{3/2}/(\sqrt{M}B))$ I/Os assuming $M \geq E^{\Omega(1)}$. Our results are based on a new color coding technique, which may be of independent interest.

Host: Rolf Fagerberg