2018 Research Interest/Project Ideas

Alex Pothen

www.cs.purdue.edu/homes/apothen/
www.csc-research.org
www.cscapes.cs.purdue.edu
==

High Performance Algorithms for Massive Graphs

Graphs model data from social networks, transportation networks, data science, machine learning, the web, graphics, power grids, neuroscience, and many other application areas. The sizes of these graphs can now approach billions or even trillions of nodes and edges. To solve computational problems on these graphs efficiently, one needs to harness the parallelism available on desktop machines as well as on extreme-scale computers available at the Department of Energy Labs. We have developed parallel graph algorithms by the paradigm of approximation: often algorithms for solving problems such as graph matching, edge covers, etc., optimally do not have sufficient parallelism, but by designing approximation algorithms that have provable worst-case performance ratios, one can design low depth (number of steps in a parallel algorithm) and linear work (total number of operations) parallel algorithms. We have worked with Intel and IBM to develop efficient software for such graph computations.

The application areas for this work are as mentioned above. We have applied matching and covering problems to the k-Nearest Neighbor (kNN) Graph computation used widely for creating sparse graph representations of dense, noisy data, and for adaptive anonymization of data to be published while respecting privacy constraints. We are also part of the U.S. Department of Energy's Exascale Computing Project (ECP) for co-designing a computer capable of 10^18 operations per second, through the ExaGraph project.