

Analyzing Massive Social Networks Using Multicore and Multithreaded Architectures

David Bader

Georgia Institute of Technology, USA

Abstract. Emerging real-world graph problems include detecting community structure in large social networks, improving the resilience of the electric power grid, and detecting and preventing disease in human populations. Unlike traditional applications in computational science and engineering, solving these problems at scale often raises new challenges because of sparsity and the lack of locality in the data, the need for additional research on scalable algorithms and development of frameworks for solving these problems on high performance computers, and the need for improved models that also capture the noise and bias inherent in the torrential data streams. The explosion of real-world graph data poses a substantial challenge: How can we analyze constantly changing graphs with billions of vertices? Our approach leverages the Cray XMT's fine-grained parallelism and flat memory model to scale to massive graphs. On the Cray XMT, our static graph characterization package GraphCT summarizes such massive graphs, and our ongoing STINGER streaming work updates clustering coefficients on massive graphs at a rate of tens of thousands updates per second.