Anomaly Detection in Very Large Graphs: Modeling and Computational Considerations

Benjamin Miller, Nicholas Arcolano, Edward Rutledge, Matthew Schmidt, and Nadya Bliss, Massachusetts Institute of Technology

Abstract
Graph theory provides an intuitive mathematical foundation for dealing with relational data, but there are numerous computational challenges in the detection of interesting behavior within small subsets of vertices, especially as the graphs grow larger and the behavior becomes more subtle. This presentation discusses computational considerations of a residuals-based subgraph detection framework, including the implications on inference with recent statistical models. We also present scaling properties, demonstrating analysis of a billion-vertex graph using commodity hardware.

* This work is sponsored by the Intelligence Advanced Research Projects Activity (IARPA) under Air Force Contract FA8721-05-C-0002. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA or the U.S. Government.