Big Data analytics for Cyber security: A case study in change detection for evolving networks

Josephine Namayanja, Akshay Grover, Jay Gholap, Vandana P. Janeja
University of Maryland, Baltimore County

Motivation
Temporally Evolving Computer Network

- Importance of Control Nodes
- Density of Information

Research Objective
To identify pertinent changes in temporally evolving computer networks by evaluating the behavior of central (key) nodes and their impact on the network over time while utilizing an efficient data processing framework.

Challenges
- Numerous nodes and edges.
- Traffic is captured at multiple time intervals.
- Understanding Holistic Change is challenging.
- Computationally costly.

Approach
Node Selection
- Hybrid Sampling
Multi-Level Change Detection
- Coin Central Nodes
Network Level Coin Change Point Detection
- Consistent and Inconsistent (CoIn) Central Nodes
- Network Level Change Points due to CoIn Central Nodes (NL-CoIn)

Big Data Framework
- Node Selection
- Multi-Level Change Detection
- Big Data Framework
- Extensive experimental and comparative results using real world internet traces.
- Validate results with real world cyber attacks.

Results
Related Publications
- J. M. Namayanja, V.P. Janeja, Change Detection in Evolving Computer Networks: Changes in Densification and Diameter Over Time, IEEE International Conference on Intelligence and Security Informatics, 2015, Baltimore MD
- J. M. Namayanja, V.P. Janeja, Change Detection In Temporally Evolving Computer Networks: A Big Data Framework, First International Workshop on High Performance Big Graph Data Management, Analysis, and Mining, co-located with the IEEE BigData 2014 in Washington DC, USA

Cluster Specifications: Maya HPCF (NSF, UMBC)
- The hardware used in the computational studies is part of the UMBC High Performance Computing Facility (HPCF)
- 240 Nodes Cluster [16 used for Hadoop]
- Each node consists of two quad-core 2.8 GHz Intel Nehalem X5560 CPUs and 24 GB memory designed for fastest number crunching and connected by a dual-data rate (DDR) InfiniBand network.