Multi-Centrality Graph Spectral Decompositions and Their Application to Cyber Intrusion Detection

Dr. Pin-Yu Chen[1]
Dr. Sutanay Choudhury[2]
Prof. Alfred Hero[1]
[1]Department of Electrical Engineering and Computer Science
University of Michigan
[2]Pacific Northwest National Laboratory

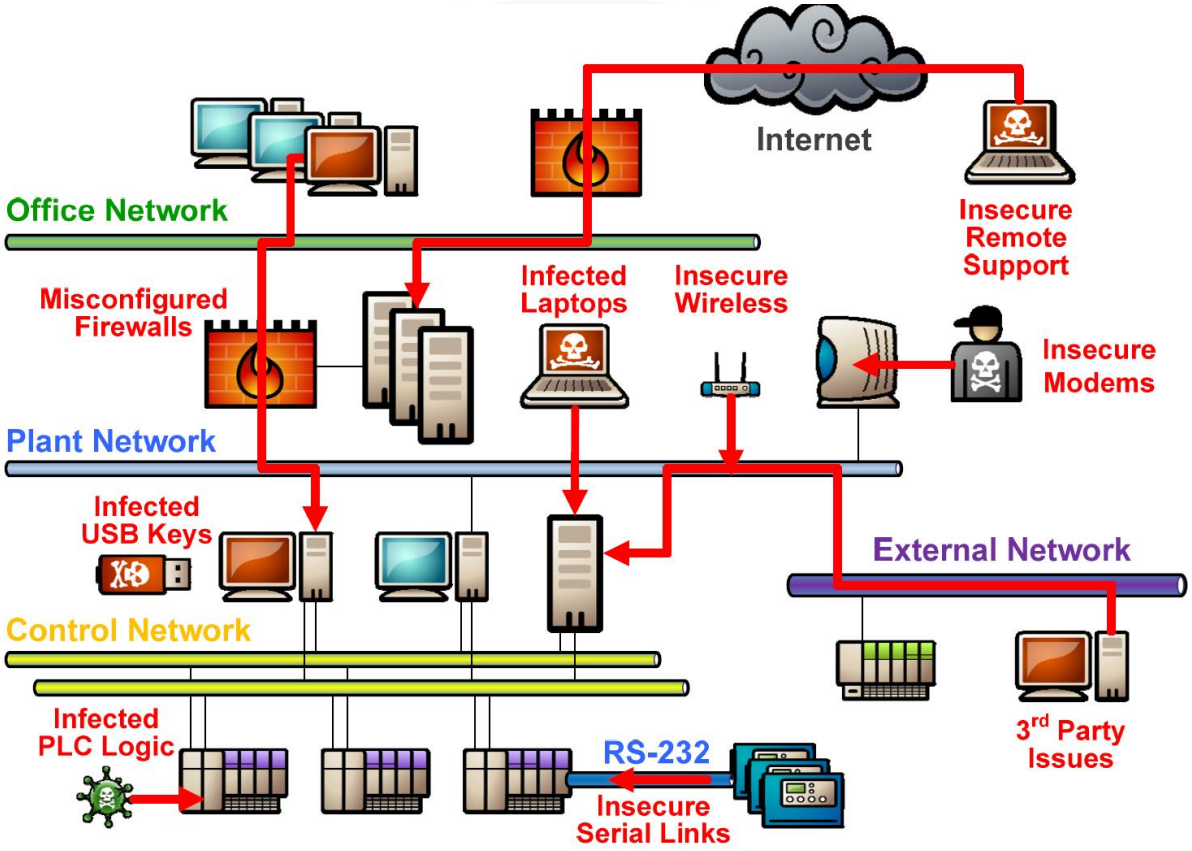
Motivation

"The IAEA has reported cases of random malware-based attacks at nuclear plants"[1]

"Cyber threats to nuclear materials, nuclear facilities and nuclear command, control and communications are becoming more sophisticated every day, and the global technical capacity to address the threat is limited."[2]

[1] Inter Press Service News Agency, Aug 17, 2015
[2] Senator Sam Nunn and Ted Turner, Nuclear Threat Initiative (NTI), 2016

Nuclear Facility Cyber Intrusion



Possible Pathways into a Control System

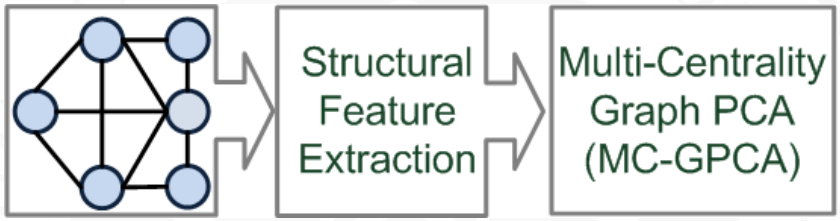
Source: Fig. 1, Eric Byres, Tofino Security White Paper, , 2012.

Cyber Security via Graph Mining

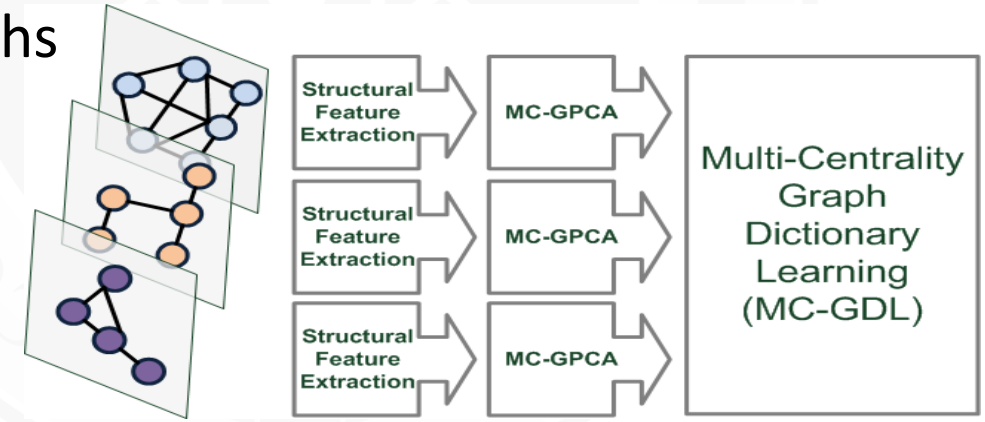
- Represent pairwise interactions between entities on network as a graph
- Extract structural features from a graph (or ensembles of graphs) for graph connectivity summarization and anomaly detection
- Transform graph representation to feature matrix representation
- Identify high vulnerability nodes
- Early detection of anomalies and attacks

Our Contributions

- Multi-centrality graph PCA (MC-GPCA) for a single graph



- Multi-centrality graph dictionary learning (MC-GDL) for ensembles of graphs



- Application to cyber intrusion detection

Ref: Chen-Choudhury-Hero, ICASSP'16

Graph Structural Feature Extraction

- **Goal:** extract structural features from a graph G of n nodes and represent them by an $n \times p$ feature matrix $\mathbf{X}$
  - p : # of extracted structural features
  - $\mathbf{X} = \left[\mathbf{x}_{\#\ \mathbf{h-hop\ walk}}\ \mathbf{x}_{\mathbf{h-hop\ weight}} \cdots \mathbf{x}_{\mathbf{deg}}\ \mathbf{x}_{\mathbf{eig}}\ \mathbf{x}_{\mathbf{betweenness}} \cdots \mathbf{x}_{\mathbf{ref}} \cdots\right]_{n \times p}$

- Three types of graph structural features are extracted:

(1) graph walk statistics

- # of h-hop walks – $\mathbf{a}^{(h+1)} = \mathbf{A}\mathbf{a}^{(h)}$
  $\mathbf{A}$: adjacency matrix of G
- total weight of h-hop walks – $\mathbf{w}^{(h+1)} = \mathbf{W}\mathbf{a}^{(h)} + \mathbf{A}\mathbf{w}^{(h)}$
  $\mathbf{W}$: edge weight matrix of G

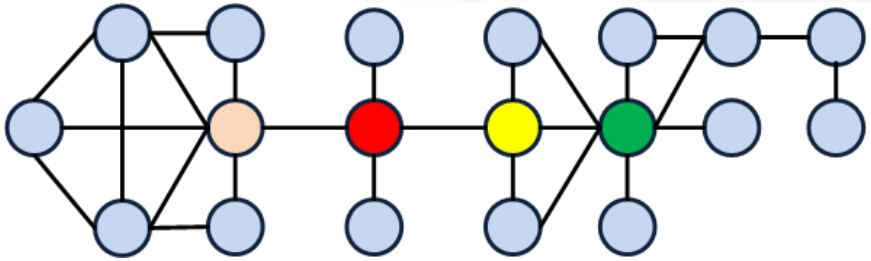
-> efficient recursive computation

Graph Structural Feature Extraction

(2) centrality measures

- One centrality measure of each node = One extracted structural feature (one column of $\mathbf{X}$)
- LFVC: local Fiedler vector centrality [Chen-Hero ICASSP'14, Comm. Mag.'14, Tran. Signal Processing'15]

Degree	• # of connections
Betweenness	• # of shortest paths
Closeness	• average hop distance
Eigenvector	• relative importance
Ego	• local betweenness
LFVC	• algebraic connectivity

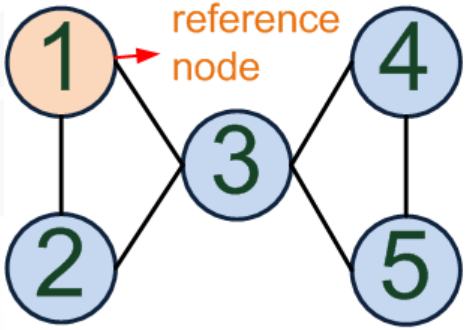


Graph Structural Feature Extraction

(3) **hop distance to a set of reference nodes**

- select some “anchor nodes” in the graph and use hop distances to these nodes as graph structural features
- enhance structural identifiability for graphs of high symmetry
- example:

  - nodes 1,2,3,4 are indistinguishable using degree feature
  - use node 1 as a reference node, then only nodes 4 and 5 are indistinguishable



$$x_{deg} = \begin{bmatrix} 2 \\ 2 \\ 4 \\ 2 \\ 2 \end{bmatrix} \quad x_{ref} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

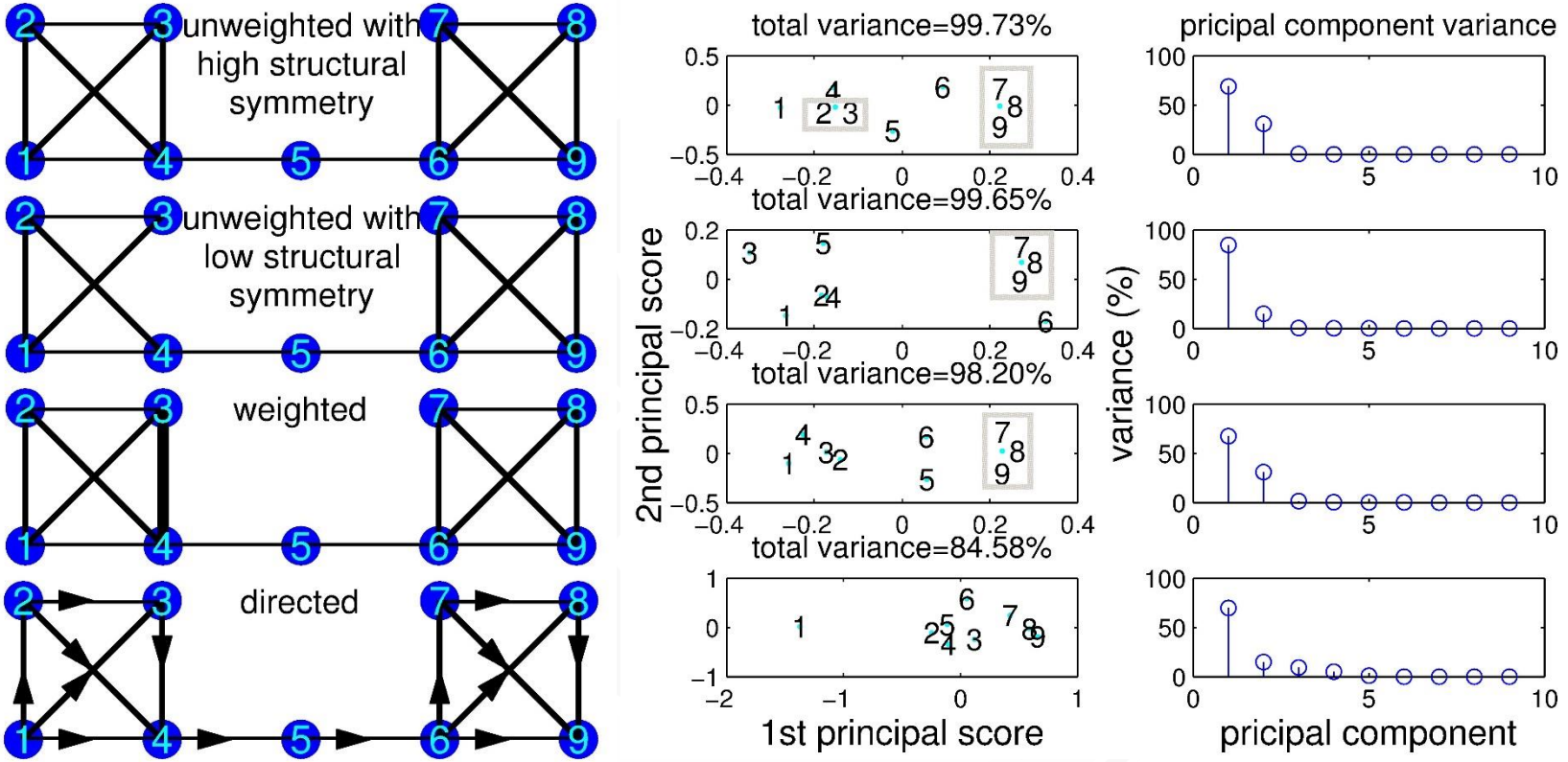
Multi-Centrality Graph PCA (MC-GPCA)

- Project (normalized and centered) structural feature matrix $\mathbf{X}$ onto a set of uncorrelated orthogonal basis
- Advantages:
  1) Decorrelation – structural features are often correlated
  2) Visualization via projection coefficients

Algorithm 1 Multi-centrality graph PCA (MC-GPCA)

Input: A graph $G = (\mathcal{V}, \mathcal{E})$, desired dimension q
Output: n structural coordinates $\mathbf{Y}$ for each node in G
1. Extract p structural vectors $\mathbf{X}$ from G
2. Normalize each column of $\mathbf{X}$ to have unit norm
3. Subtract row-wise empirical average from $\mathbf{X}$
4. Solve the right singular vectors $\mathbf{V}_q$ of $\mathbf{X}$
5. $\mathbf{Y} = \mathbf{X}\mathbf{V}_q$

MC-GPCA Illustrating Examples



extract 1-4 hop graph walk statistics, reference node = node 1, $p = 9$

Structural Difference Score (SDS)

- $\mathbf{Y}$: an $n \times q$ matrix from MC-GPCA
- SDS of node i :

$$\mathrm{SDS}(i) = \frac{\sum_{j \in \mathrm{Neighbor}(i)} \left\| row_i(\mathbf{Y}) - row_j(\mathbf{Y}) \right\|^2}{d_i + 1}$$

d_i: degree of node i

➢ Interpretation: an aggregated centrality measure of structural difference of a node and its neighboring nodes

Multi-Centrality Graph Dictionary Learning (MC-GDL)

- Given a set of g graphs. For each graph, run MC-GPCA, compute SDS, extract top z nodes of highest SDS as one column of $\mathbf{Z}$
- Sparse dictionary learning model: $\mathbf{Z} = \mathbf{DC} + \mathbf{noise}$. $\mathbf{C}$ column sparse
- K-SVD [Aharon-Elad-Bruckstein, Tran. Signal Processing'06]:

$$\mathbf{min}_{\mathbf{D},\mathbf{C}} \|\mathbf{Z} - \mathbf{DC}\|_{\mathbf{F}}^{\mathbf{2}} \text{ subject to } \left\|\mathbf{col}_{\mathbf{j}}(\mathbf{C})\right\|_{\mathbf{0}} \leq S, \forall \text{ j}$$

Algorithm 2 Multi-centrality graph dictionary learning (MC-GDL)

Input: A set of graphs $\{G_\ell\}_{\ell=1}^{g}$, number of atoms K, sparsity constraint S, number of highest SDS feature z

Output: graph structure dictionary $\mathbf{D}$, coefficient matrix $\mathbf{C}$

1. Obtain z highest SDS for each graph as columns of $\mathbf{Z}$
2. Subtract column-wise empirical average from $\mathbf{Z}$
3. Perform K-SVD on $\mathbf{Z}$ to obtain $\mathbf{D}$ and $\mathbf{C}$

Application: Cyber Intrusion Detection

- The University of New Brunswick (UNB) intrusion detection dataset - 7 days of graph connectivity patterns of a cyber system [3]

Dataset	# nodes	# edges	Description
Day 1	5357	12887	Normal activity
Day 2	2631	5614	Normal activity
Day 3	3052	5406	Infiltrating attack and normal activity
Day 4	8221	12594	HTTP denial of service attack and normal activity
Day 5	24062	32848	Distributed denial of service attack using Botnet
Day 6	5638	13958	Normal activity
Day 7	4738	11492	Brute force SSH attack and normal activity

Extracted structural features for each graph:

1) 1-20 hop graph walk statistics
2) 6 centrality measures
3) 10 reference nodes of highest degree

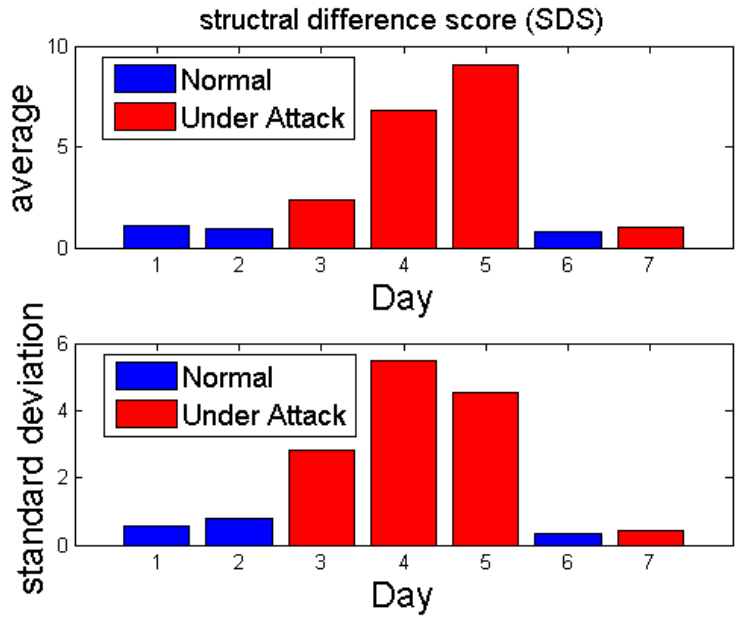
In total p = 56 features

Ref: Chen-Choudhury-Hero, ICASSP'16

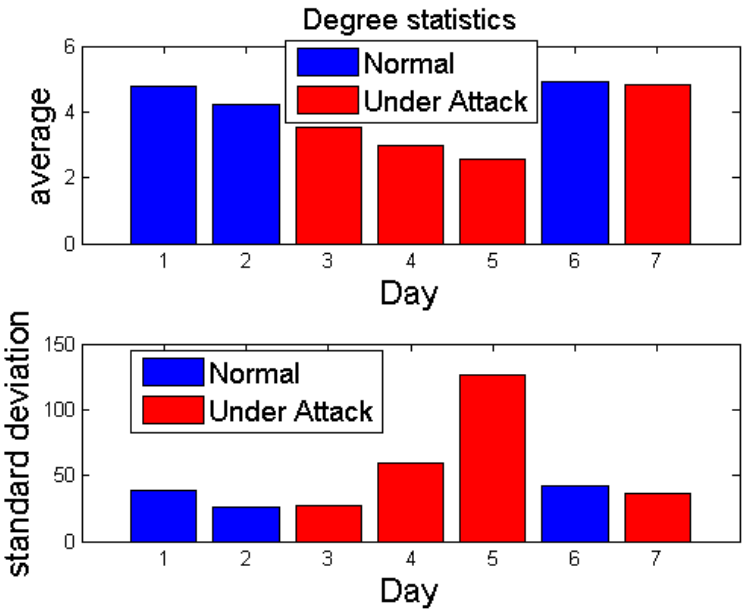
[3] Shiravi-Shiravi-Tavallaee-Ghorbani, *Computers & Security*'12

Performance Evaluation - MC-GPCA

SDS score (proposed)



Degree score

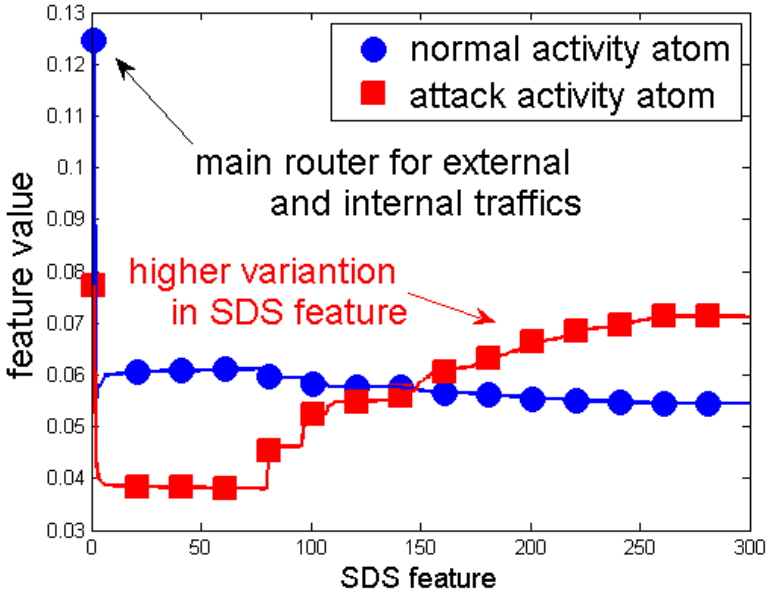


✓ The proposed SDS from MC-GPCA is an effective indicator of intrusions resulting in anomalous connectivity patterns (Days 3,4,5)

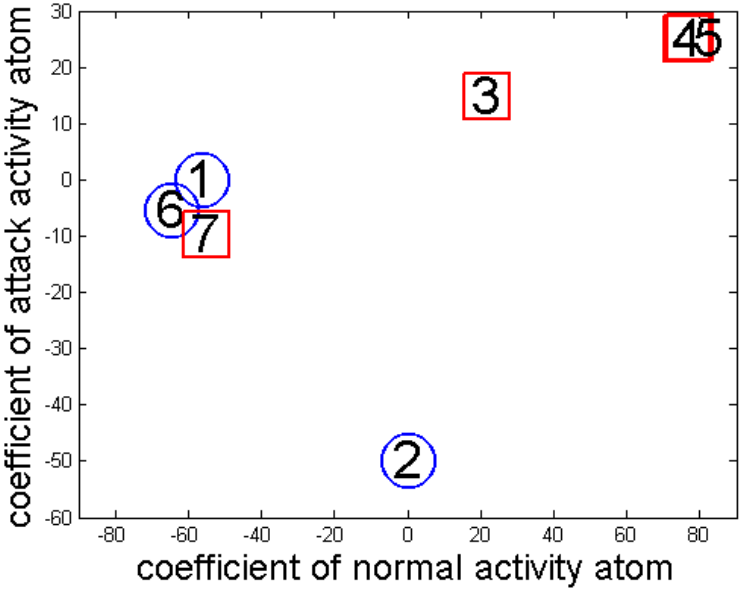
Ref: Chen-Choudhury-Hero, ICASSP'16

Performance Evaluation - MC-GDL

Atoms from MC-GDL



Coefficients from MC-GDL



- ✓ The atoms learned from MC-GDL reflect normal and anomalous connectivity patterns, and the coefficient matrix **C** from MC-GDL can be used for attack classification.

Ref: Chen-Choudhury-Hero, ICASSP'16

Conclusion

- Nuclear facilities are vulnerable to cyber attacks
- Tools for early detection of cyberattacks are essential
- Proposed a multi-centrality decomposition of at-risk networks (MC-GPCA & MC-GDL)
- Proposed a structural difference score (SDS) for detecting anomalous connectivity patterns
- Demonstrated on cyber intrusion database based on pairwise connectivity of nodes

Acknowledgements

- This work was partially supported by the Consortium for Verification Technology under Department of Energy National Nuclear Security Administration award number DE-NA0002534 and by the Asymmetric Resilient Cyber Security initiative at Pacific Northwest National Laboratory.

- Contact Info: {pinyu,hero}@umich.edu

Ref: Chen-Choudhury-Hero, ICASSP'16