Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Cyber Big-Data Analytics

GEORGE CHIN, SUTANAY CHOUDHURY, KHUSHBU AGARWAL, VIVEK DATLA, PATRICK PAULSON, DENNIS THOMAS, PAK WONG, SETH THOMPSON, NATE KRUSSELL

Pacific Northwest National Laboratory

September 30, 2015

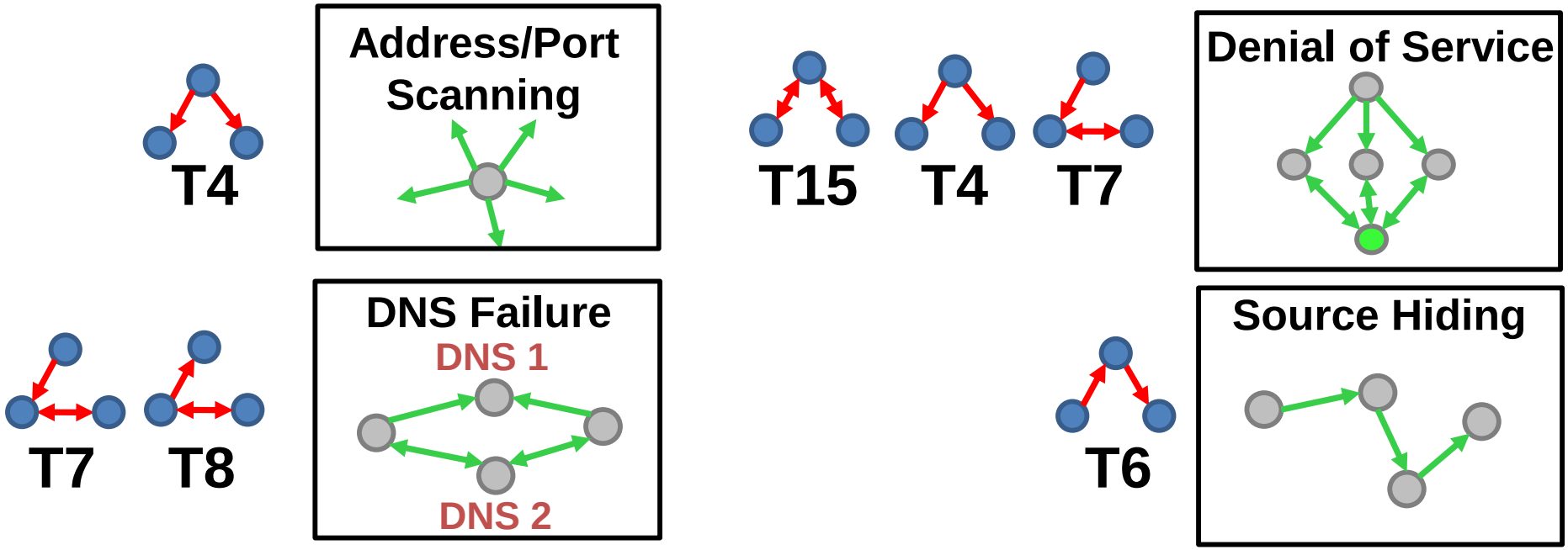
PNNL-SA-113259

Idaho Bailiff Cyberanalytics Visualization Examples

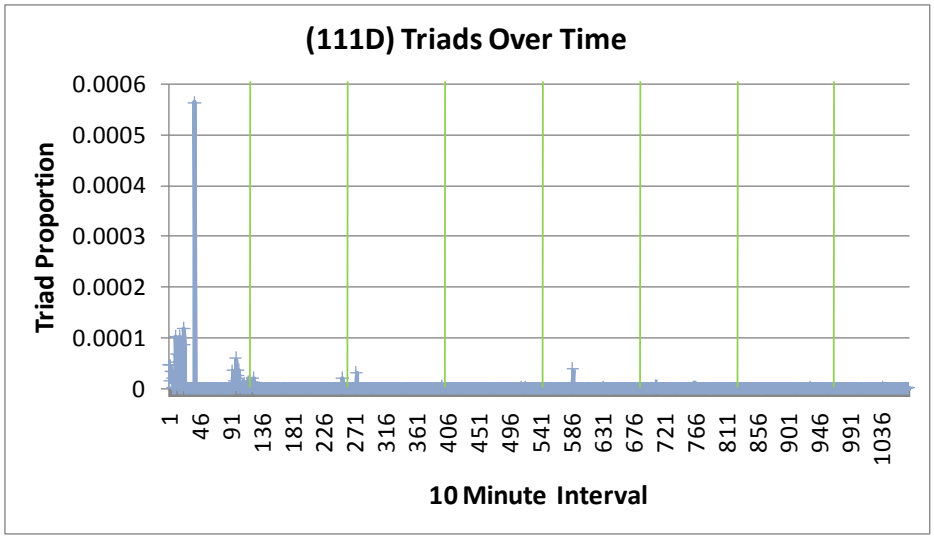
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- Idaho Bailiff project focuses on following areas:
  - High performance computing platforms
  - Graph algorithms
  - Streaming data
  - Big data
  - Cyberanalytics applications
- Visualizations developed have following properties
  - Network/graph visualizations
  - Secondary area of research and development
  - Support pattern discovery
  - Extend to big data and/or fast data streaming rates

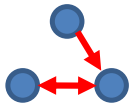
Computer Network Threat Detection Using Dynamic Triadic Analysis

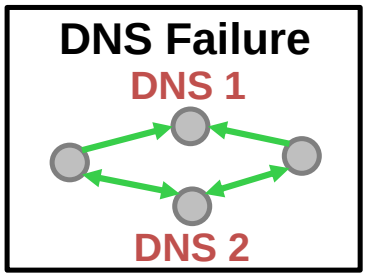


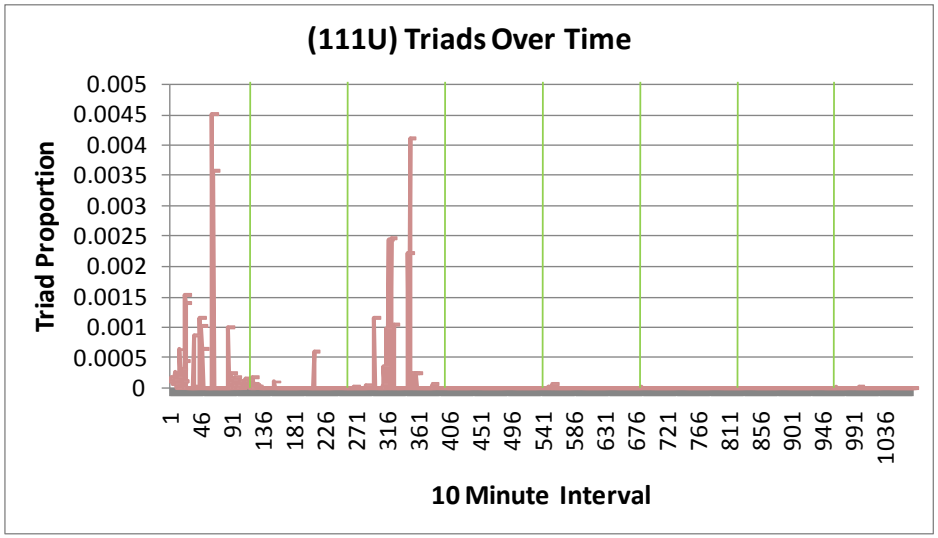
Computer Network Threat Detection Using Dynamic Triadic Analysis



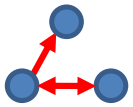
All initiator triads (T7)







All timeout triads (T8)



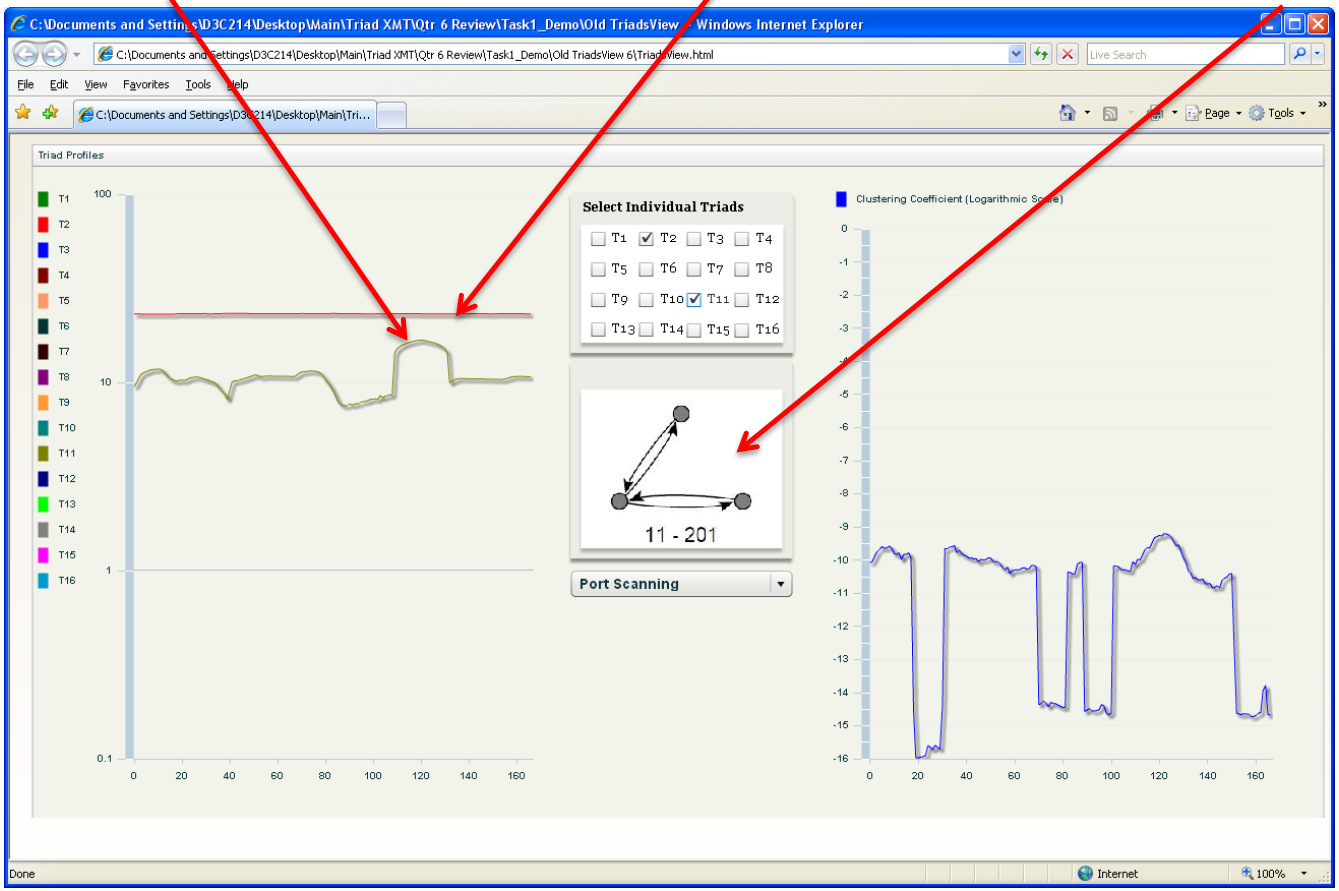
Computer Network Threat Detection Using Dynamic Triadic Analysis

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Changes in Specific Triad Patterns Signify a Potential Network Threat

While Other Triad Patterns are Unaffected

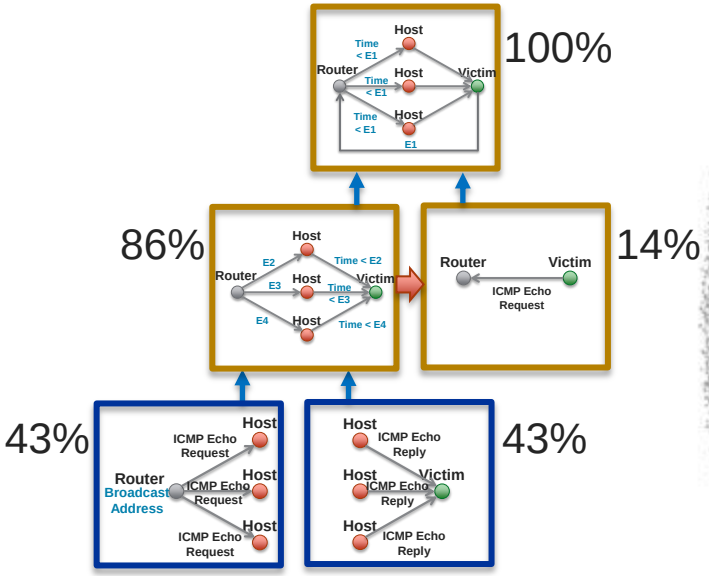
Specific Triad Types Indicate Specific Threats

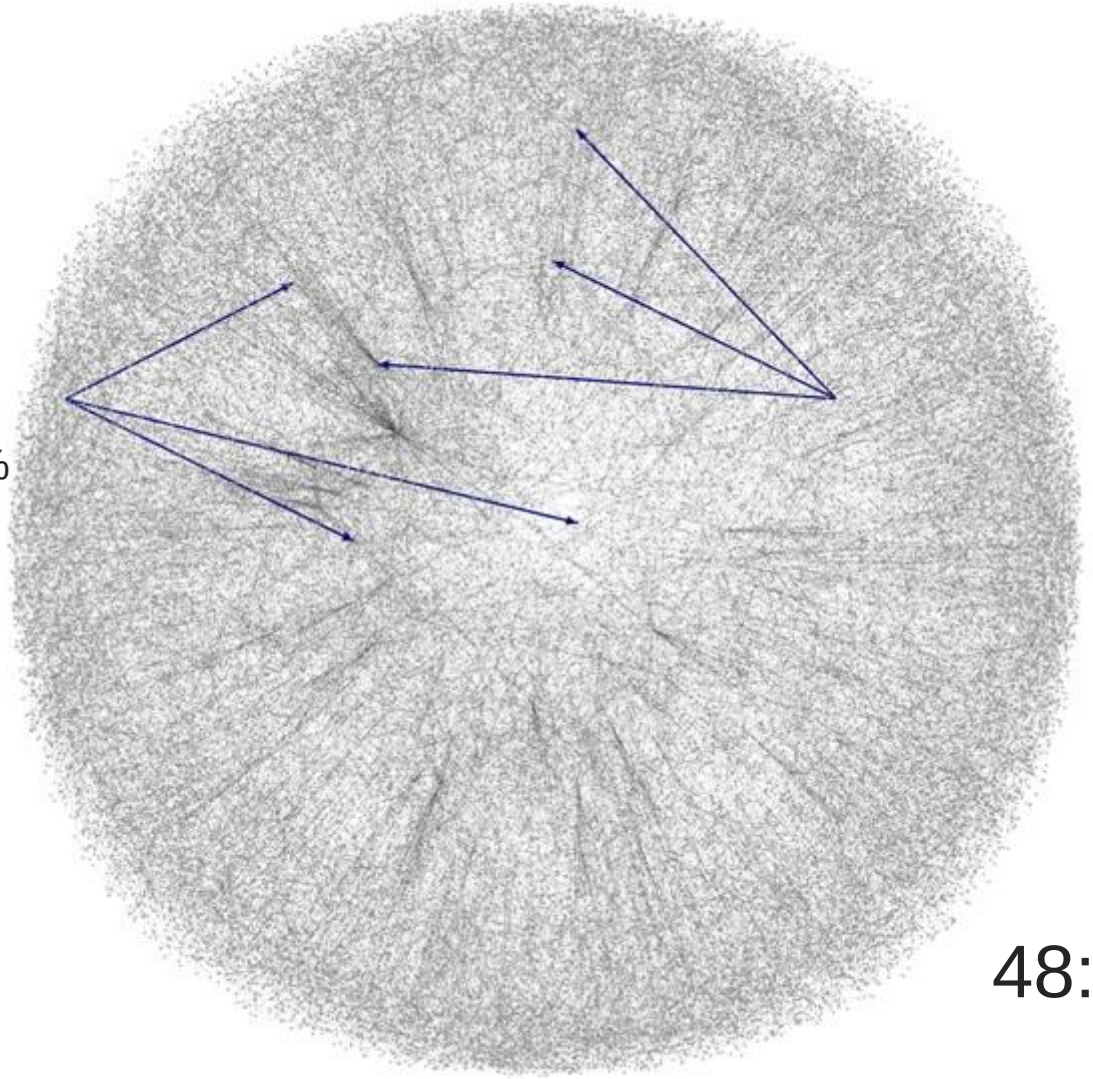


DDoS Attack Query

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Subgraph Join Tree



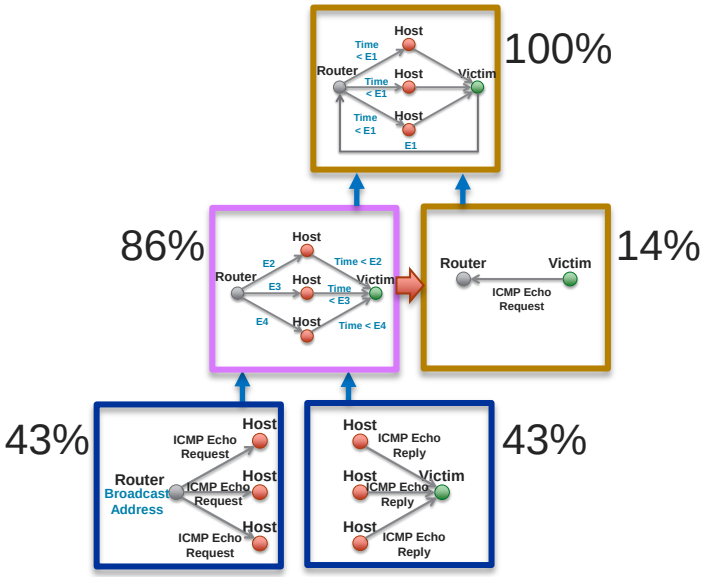


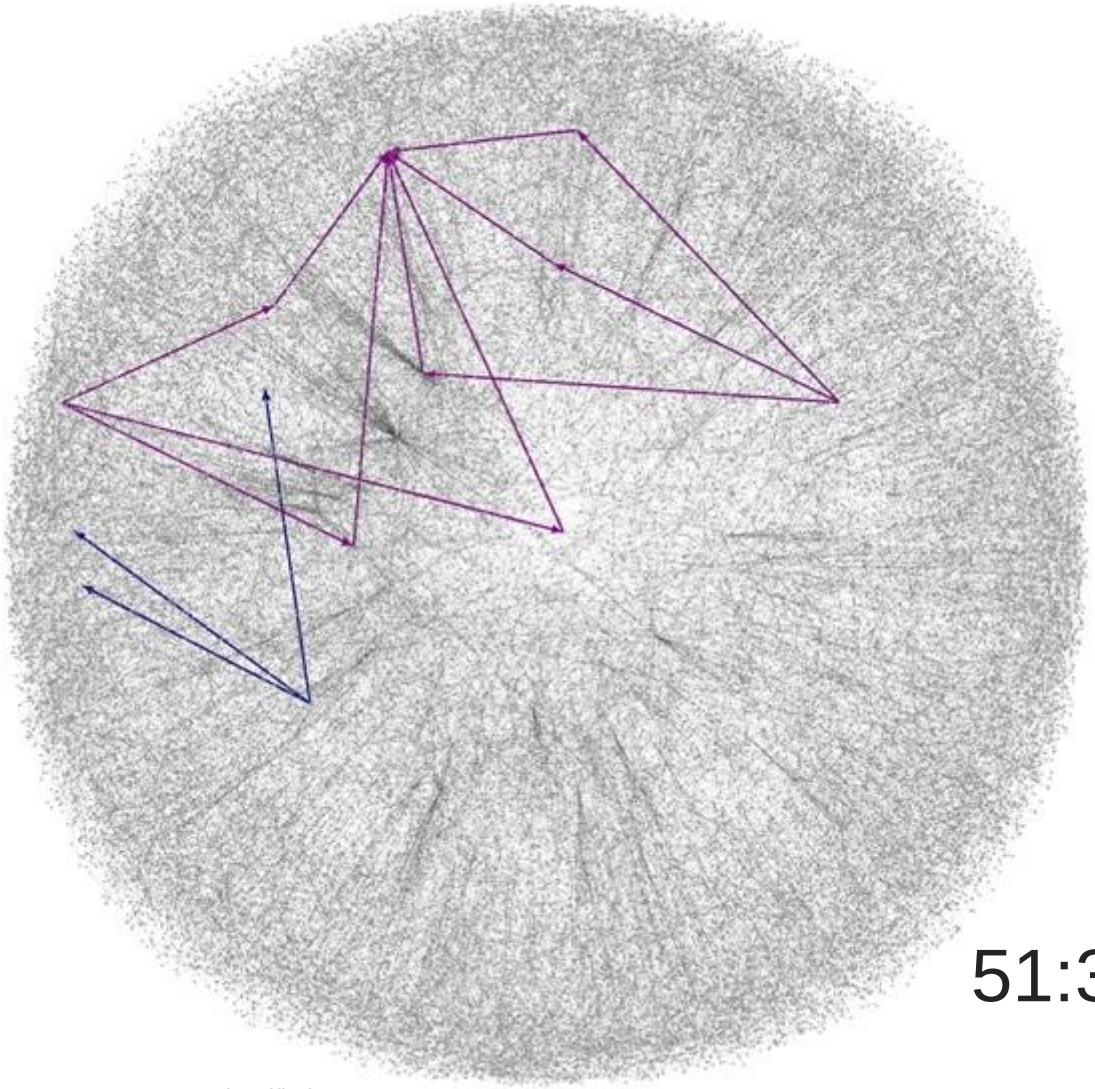
48:06

Emerging Cyberattack Patterns

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Subgraph Join Tree



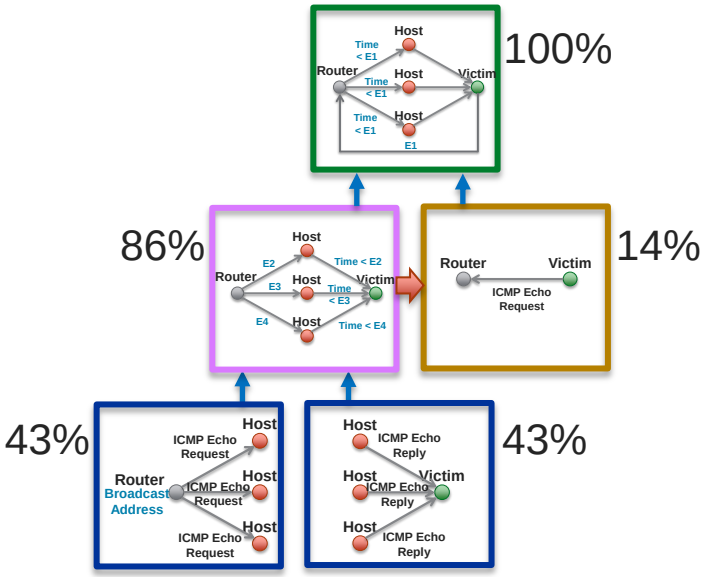


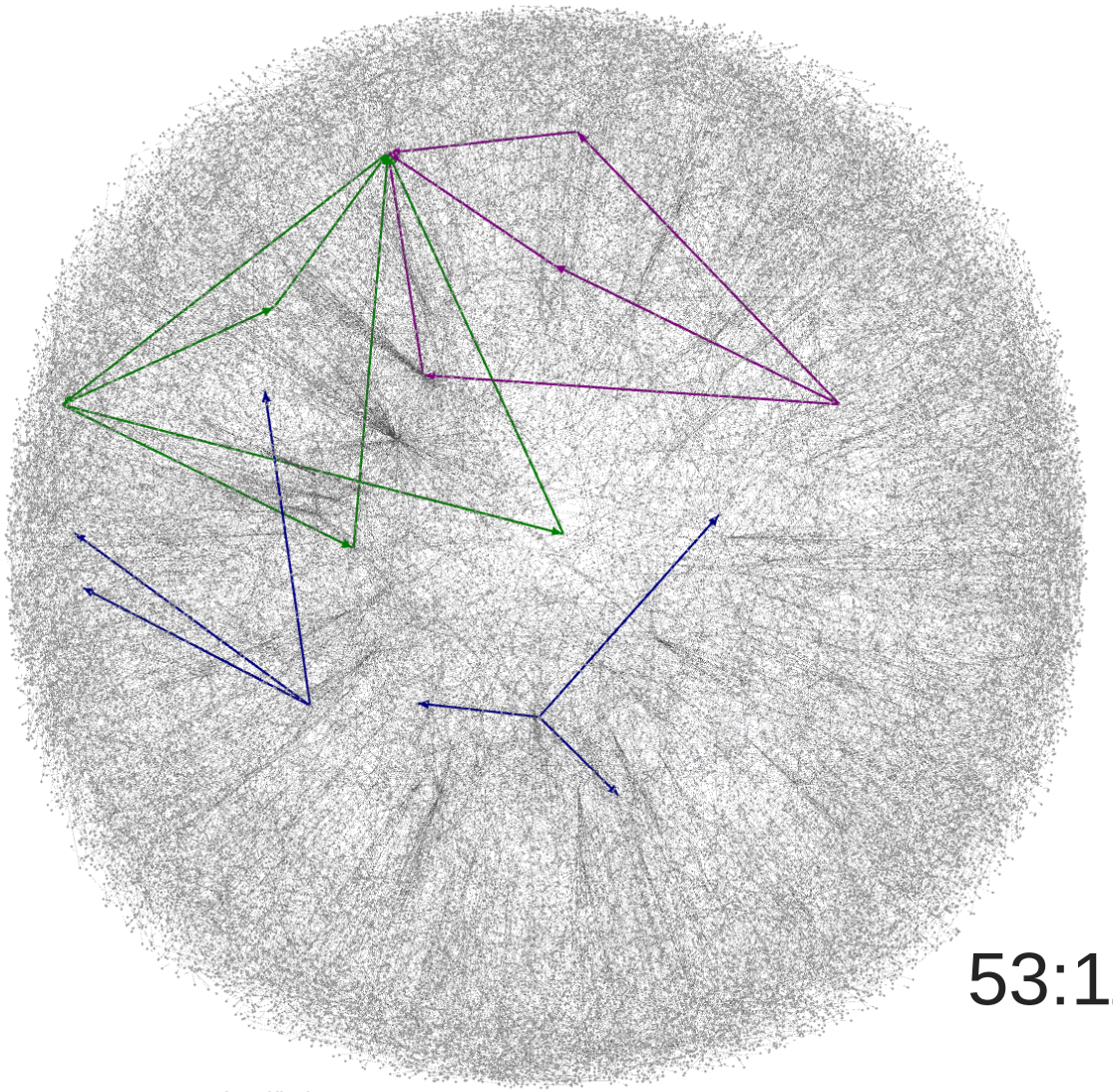
51:39

Emerging Cyberattack Patterns

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Subgraph Join Tree

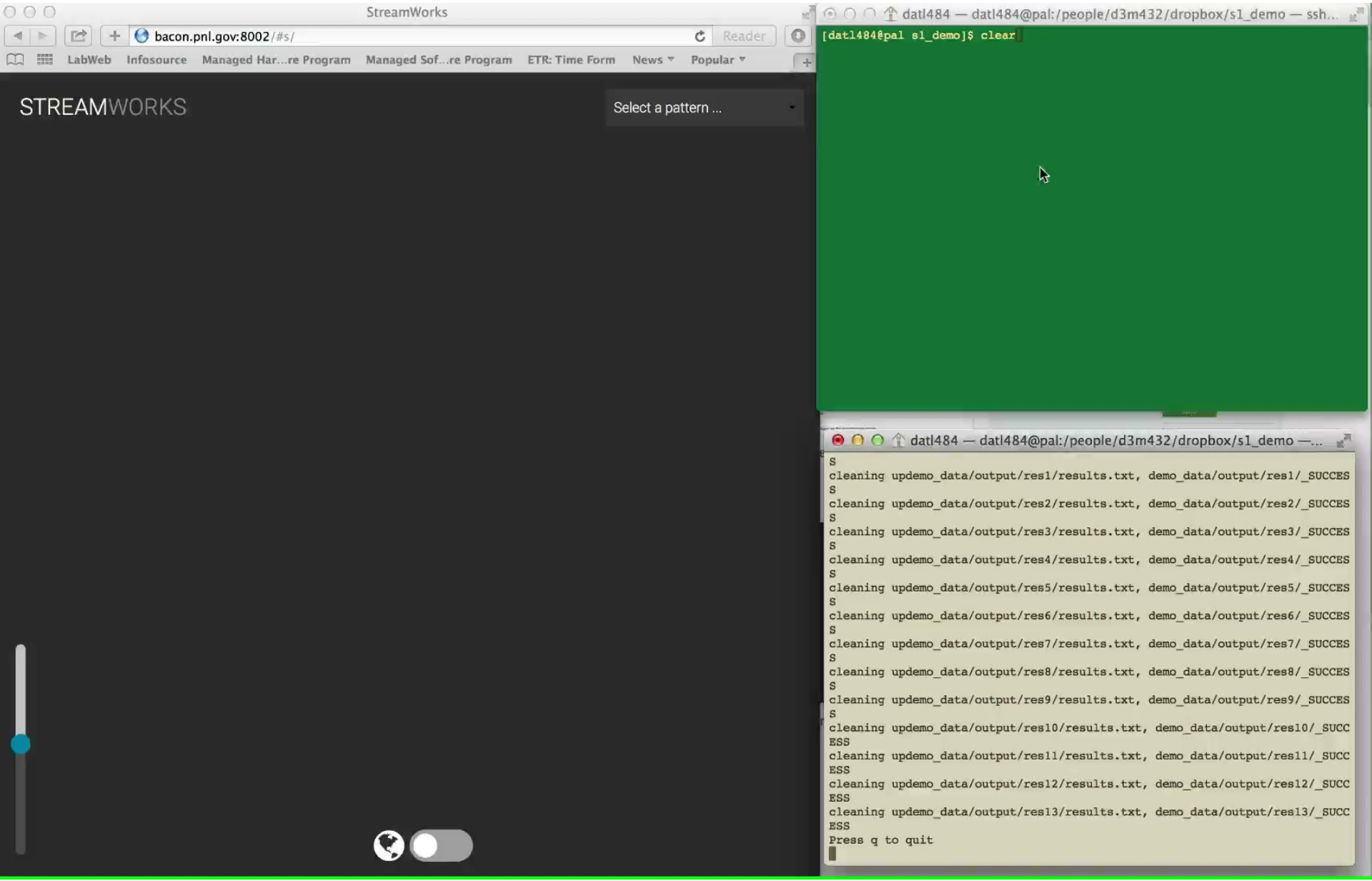




53:11

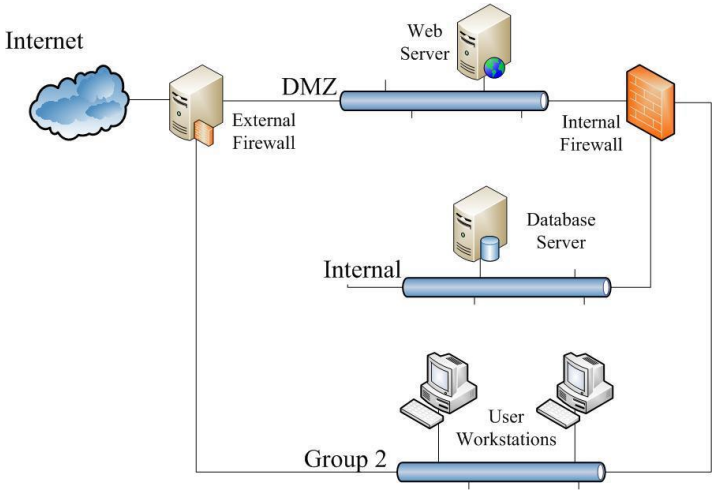
Emerging Cyberattack Patterns Video

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

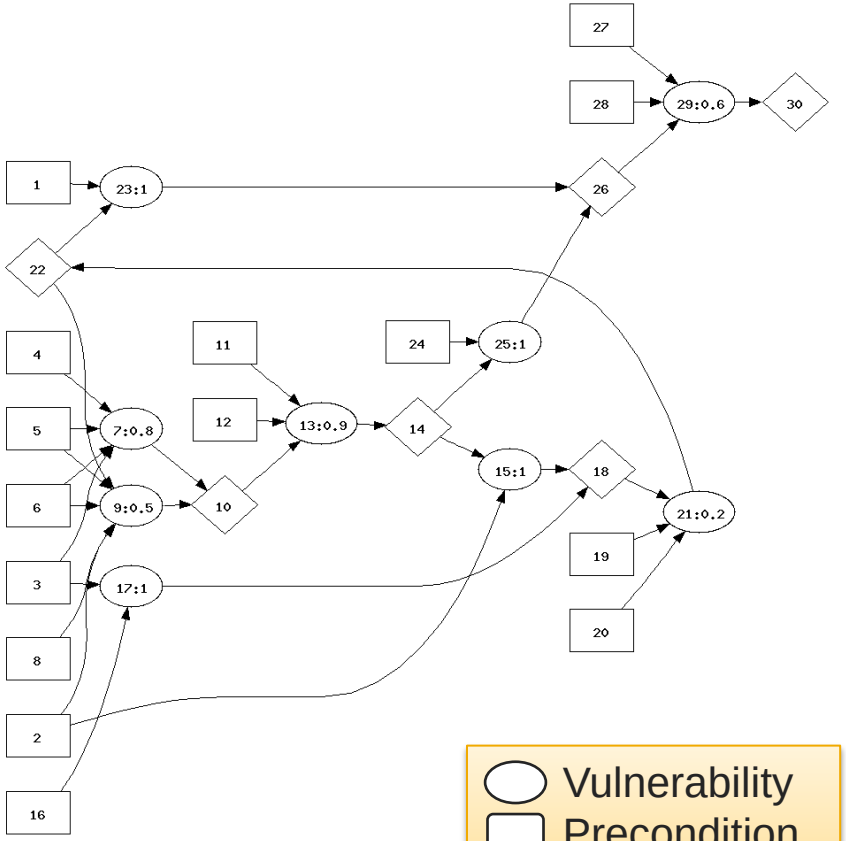


Probabilistic Vulnerability Graph

Network Configuration

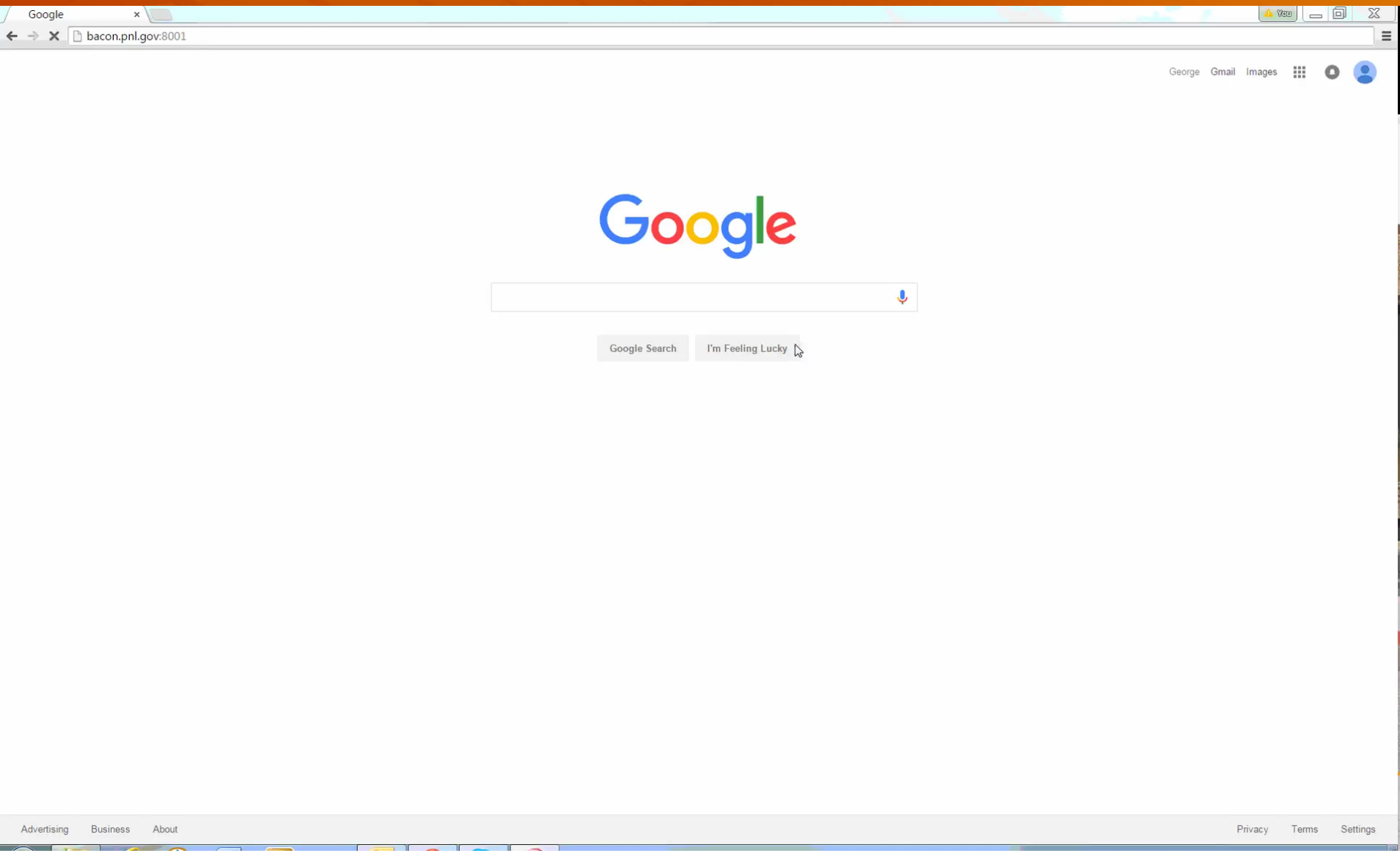


Vulnerability Graph



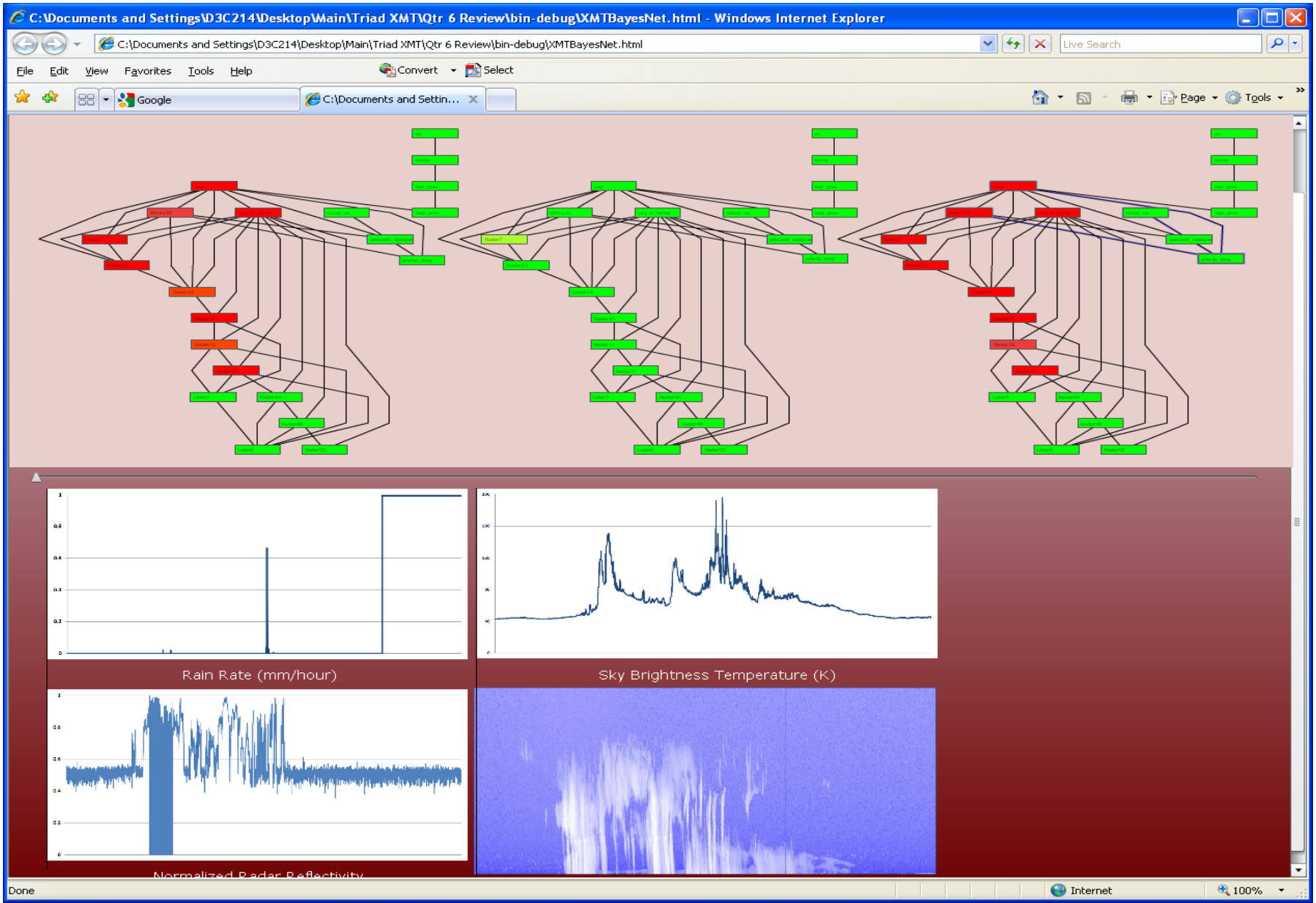
Probabilistic Vulnerability Graph Video

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*



Atmospheric Sensor Network Dynamic Bayesian Network

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*



Applications to the Future Power Grid

- A graph or network is a natural and intuitive way to visually represent the power grid
- Like computer network flow data, power grid data may be streaming in at fast rates
- Pattern discovery and network analysis approaches algorithms may be useful in representing and detecting anomalies and behaviors in the power grid
- Probabilistic graphical models may also be useful in power grid predictions and diagnostics
- Associated network visualizations should be equally enabling and useful