SAND2019-7242PE

5.3.3 Autonomous Malware Detection, Isolation and Mitigation

GMI Topic Area 5.3.3

DOE 2019 Grid Modernization Lab Call

Team: James Obert (SNL), Adrian Chavez (SNL), Chris Lamb (SNL), Jovana Helms (LLNL), Tony Markel (NREL), James Boston (CPS Energy)

Sandia National Labs

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Date 6/20/2019

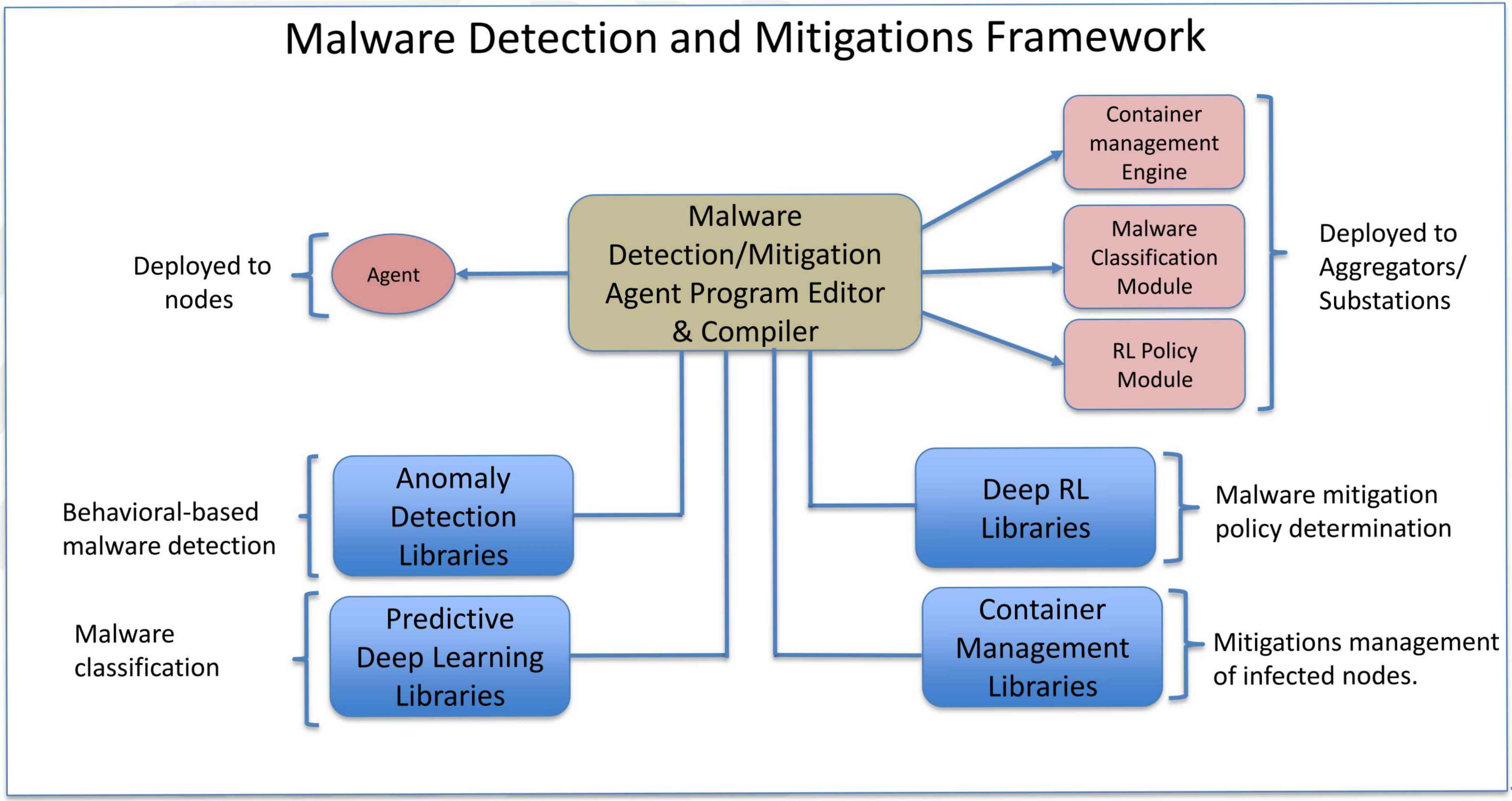
Objectives

- **Overview and Problem Statement**
  - ✓ Malware protection gaps widely exist in grid distribution nodes due to the shear number, types and distributed nature of nodes.
  - ✓ Vulnerabilities in distribution nodes open the door to potential infection of connected bulk power control and data acquisition nodes.
  - ✓ Malware infected node recovery mechanisms either do not exist or are not currently adequate.
  - ✓ This proposed work builds upon work done on GMLC project 1.4.23.
- **Objectives of Project**
  - ✓ Research and develop machine (ML), deep (DL) and reinforcement learning (RL) algorithms and framework for malware detection and mitigation.
  - ✓ Develop a framework capable of creating executables that enable power grid distribution and control nodes to autonomously detect, isolate and remove malware.
  - ✓ Framework executables restore the infected node's OS image and applications.

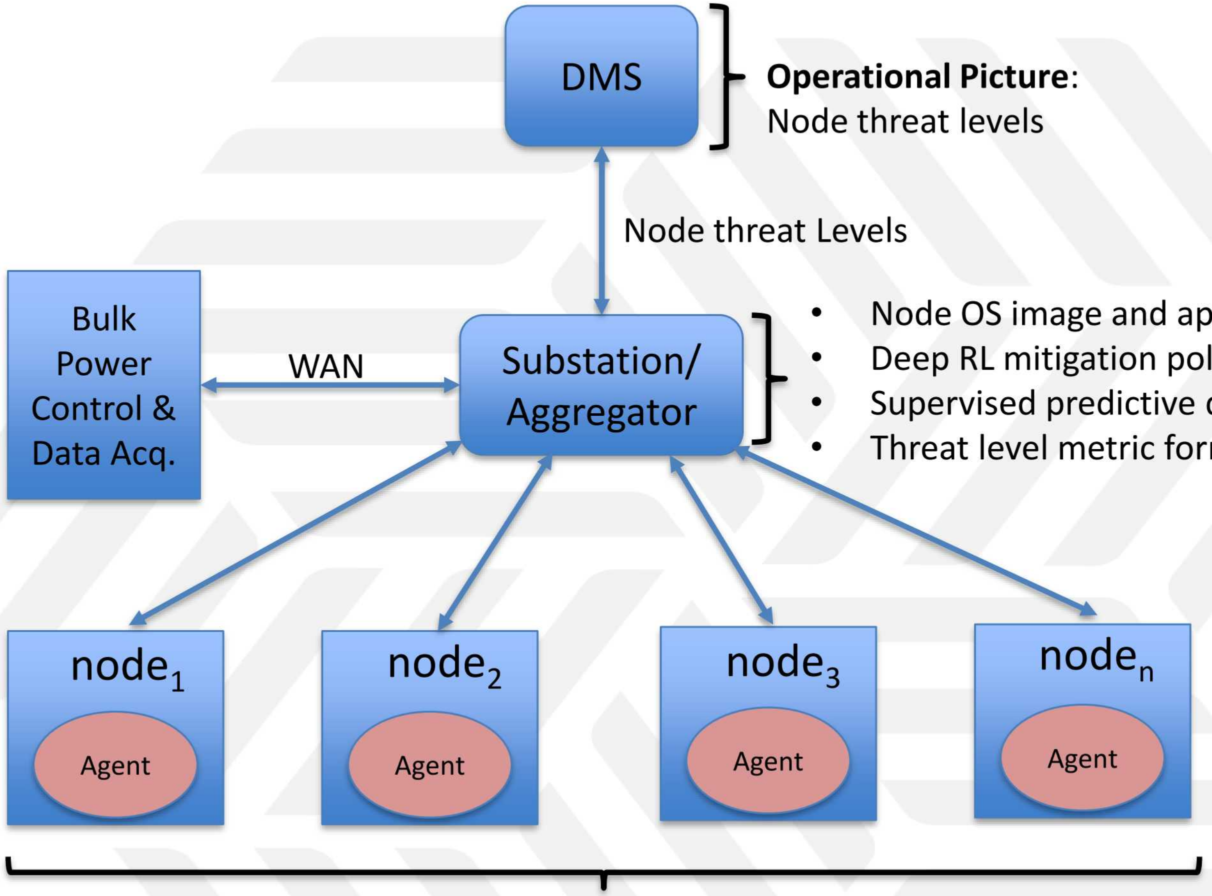
Innovation and Impact

- **Innovations:**
  - ✓ Framework enables modeling of new and existing grid malware threats.
  - ✓ Framework creates light-weight executables via a chosen integrated development environment (IDE) for deployment to operational grid OT and IT nodes.
  - ✓ Deployed executables capable of malware threat detection, isolation and recovery of malware infected nodes.
- **Impact:**
  - ✓ Wide deployment of light-weight malware detection/mitigation executable to grid distribution and endpoint nodes where it has not formerly been feasible (developed within 18-24 month time period).
  - ✓ Reduces malware attack surface at grid's vulnerable edge.
  - ✓ Two-tiered detection scheme increases accuracy of malware detection.
  - ✓ Node threat level communications back to DMS enables grid operational view and timely strategic response.
- **Support of Topic Area Objectives :**
  - ✓ Overall, the proposed technology will provide system-wide cyber-resilience in the face of malware intended to disrupt energy delivery in nodes deployed in energy sector OT and IT domains.

Technical Approach



Technical Approach

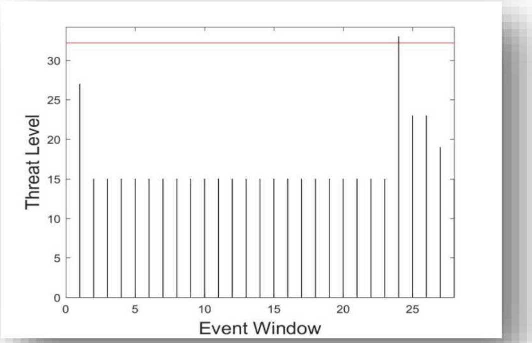


- Node OS image and app container management.
- Deep RL mitigation policy-engine.
- Supervised predictive deep learning.
- Threat level metric formation.

- Light-weight RL mitigations policy engine.
- Software profiler for dynamic behavioral-based malware anomaly detection.
- Application container managed by container manager.
- Light-weight unsupervised learning algorithms for malware anomaly detection.

Path threat Levels are composed of substation threat metric (derived in GMLC project 1.4.23) and node malware threat level:

T = malware threat metric + substation threat metric.



Technical Approach

- **Summary of Key High Level Tasks**
  - ✓ Based on surveyed grid malware threats, research and develop a malware detection and mitigation framework that enables modeling of new and existing grid malware threats and mitigation techniques.
  - ✓ Research and evaluate malware threat and mitigation executables deployed within grid distribution system nodes.

- **Key Activities, Time Periods & Highlighted Tasks Delivered within (18-24 mo.)**
  - ✓ Gather relevant malware and baseline dataset for training of detection and mitigation algorithms. (FY20)
  - ✓ Research and develop ML, DL and RL algorithms. (FY20 – FY22)
  - ✓ Development of extensible software framework and incorporation of algorithms into the framework. (FY22)
  - ✓ Validation and testing of framework and deployable executables with CPS Energy test labs, and the ability for the executables to relay threat level status to the NREL ADMS platform. (FY22)

Technical Approach

- **Leveraging several existing tools and libraries including:**
  - ✓ Malware detection model training
    - SNL's Forensics Analysis Repository for Malware (FARM)
    - VirusTotal malware repository
    - Contagio malware repository
  - ✓ Malware classification
    - SNL's Avatar ensemble learning libraries
  - ✓ Malware mitigation actions
    - CEDS CAPSec grid container management
  - ✓ Binary translator for static code analysis
    - LLNL's ROSE compiler/translator.
  - ✓ Cyber-physical operational picture
    - NREL's Advanced Distribution Management System, ADMS
  - ✓ Malware detection modeling and mitigations policy formation
    - Open source: Pytorch, DL4J, RL4J.

Team and Resources

Laboratories

- Sandia National Labs
  - James Obert, jobert@sandia.gov
  - Adrian Chavez, adrchav@sandia.gov
  - Chris Lamb, cclamb@sandia.gov
- Lawrence Livermore National Labs
  - Jovana Helms, helms7@llnl.gov
- NREL
  - Tony Markel, Tony.Markel@nrel.gov

Partners

- CPS Energy (pilot field testing)
- ARM Ecosystems (end node adaptations)
- Schweitzer Engineering Labs (substation agents)
- NERC – EISAC (metrics, methods & tools)
- Georgia Tech (control theory consulting)

Proposed Budget Range:

$4.5 - 5 million over three years

Connection to 2019 Grid Mod Lab Call principles, GMLC MYPP and GMLC Projects

- Project 1.4.23: Threat Detection and Response with Data Analytics.
  - ✓ Leverage the data analytics and threat detection knowledge gained from this project in researching and developing grid node threat level detection algorithms.
  - ✓ Leverage the threat response knowledge gained from this project in developing malware mitigation algorithms.
- Integrated Multi Scale Data Analytics and Machine Learning for the Grid.
  - ✓ Leverage the data analytics and machine learning knowledge gained from this project in researching and developing specific at scale malware detection algorithms.

Concept Maturity and Risk

- Determine what knowledge gaps exist in concept development between now and final August presentation that reflect challenges or areas of risk:
  - ✓ Cost share arrangements with utility partner.
  - ✓ Test facility usage plans with utility partner.
  - ✓ Technology transfer plans with commercial partners.

- Identify questions for DOE regarding issues facing the project or alternatives considered
  - ✓ Will it be possible in the future to use the developed malware framework to meet the requirements of other GMI projects?

Questions?

Backup slides follow

Determining Threat Level

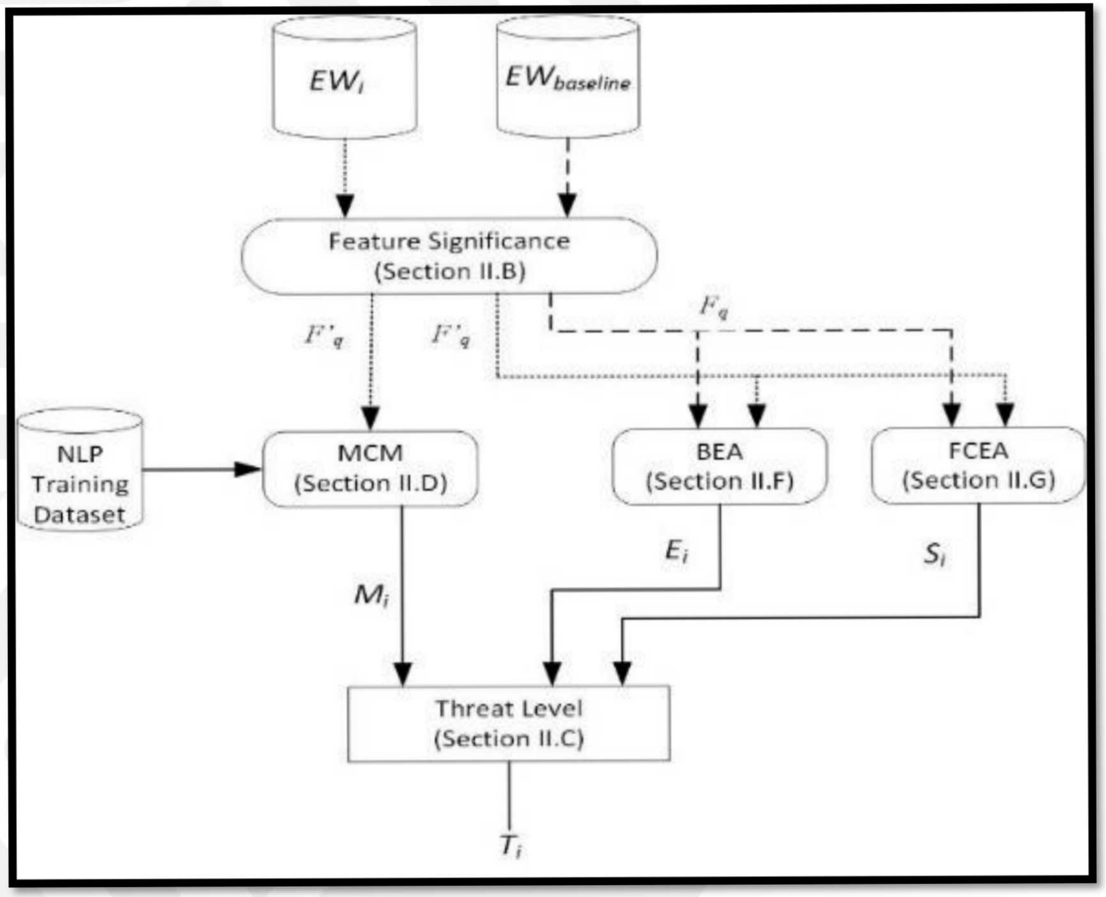


Figure 1: Determination of substation threat level

Threat Level Determination

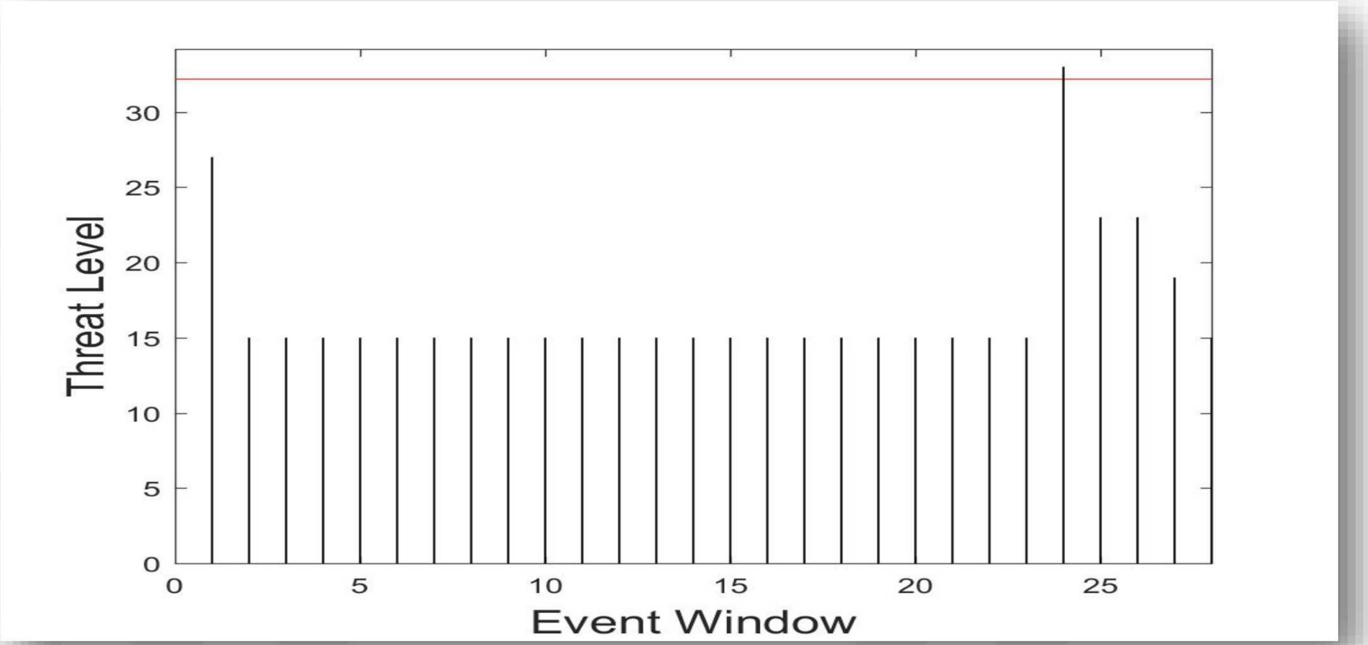


Figure 2: Threat Level Detection

M_i : *message class metric*
E_i : *Critical Feature Entropy Changes*
S_i : *Signature Match Likelihood*
C_i : *Cyber-Physical metric*
P_i : *Physical Security metric*
T_i : *Threat Level* ; ω_n : *Weight*

$$T_i = \omega_1 M_i + \omega_2 E_i + \omega_3 S_i + \omega_4 C_i + \omega_5 P_i$$