Machine Learning for Information Assurance in SCADA Systems

Richard Alcalde, Peter Bayiokos, Constanza Cabrera-Mendoza, Sabrin Kaur Guron, Wildenslo Osias, Charles C. Tappert, Avery Leider, and Krishna Bathula

Seidenberg School of Computer Science and Information Systems
Pace University

What is a SCADA System?

- ➔ Supervisory Control and Data Acquisition Systems
- ➔ Field: Industry Setting
  - ◆ i.e. power plants, manufacturing and assembly lines, chemical plants, water supply networks
- ➔ Use: Remotely control industrial machines

What is the DANGER?: Technologically dependent systems have high levels of risk if a threat finds a vulnerability. This has a tremendous impact on industries infrastructure and clients.

How can we help?

Automation of Cyber-Intrusion Classifications through the use of Machine Learning

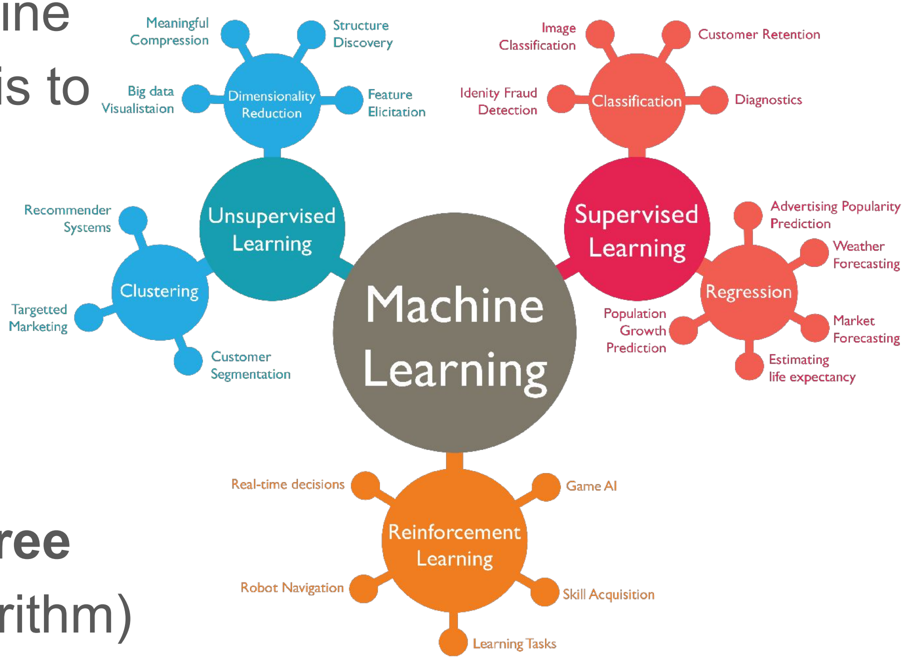
Objective: Create a Proof of Concept that demonstrates this is achievable through research of current and past experiments within the field, and implementation of that which we learn

Machine Learning: A Quick Review

ML a variety of methods to teach a machine to recognize specific patterns and use this to identify and classify data

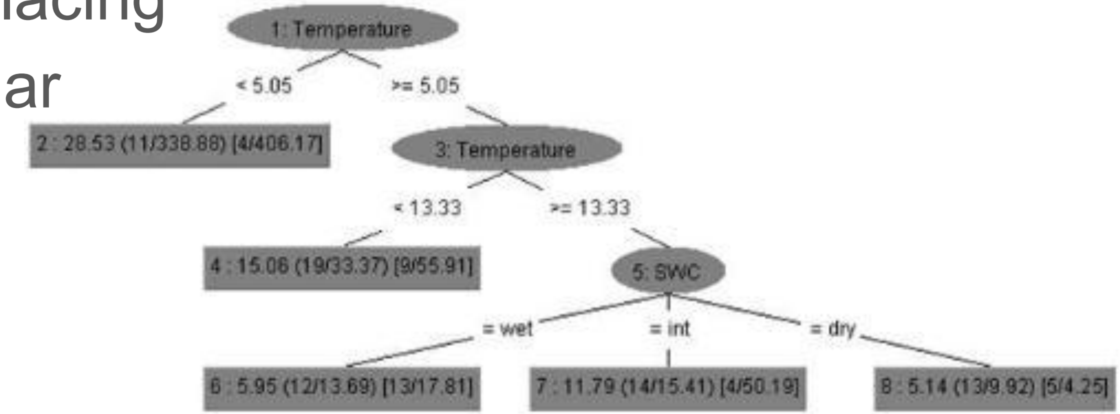
- ➔ Unsupervised Learning
- ➔ Supervised Learning
- ➔ Reinforcement Learning

We selected: **Reduced Error Pruning Tree** (Classified as a supervised learning algorithm)



Reduced Error Pruning Tree

- Decision making tree algorithm
- Based off of C4.5, data mining algorithm for data classifying.
- Reduces errors by replacing nodes with most popular classes



Literature Review Performed

- ➔ Security Issues in SCADA Networks, (2006)
- ➔ Sustainable Security for Infrastructure SCADA, (2003)
- ➔ Guide to Industrial Control Systems (ICS) Security, (2011)
- ➔ A Survey of Approaches Combining Safety and Security for Industrial Control Systems, (2015)

MITRE ATT&CK for Industrial Control Systems Framework

- ➔ Knowledge base to determine actions taken by attackers
- ➔ Describes tactics, techniques, software, and groups
- ➔ Use framework to gain clarity on attack type, severity level, and remediation steps.

Implementing the Information Learned

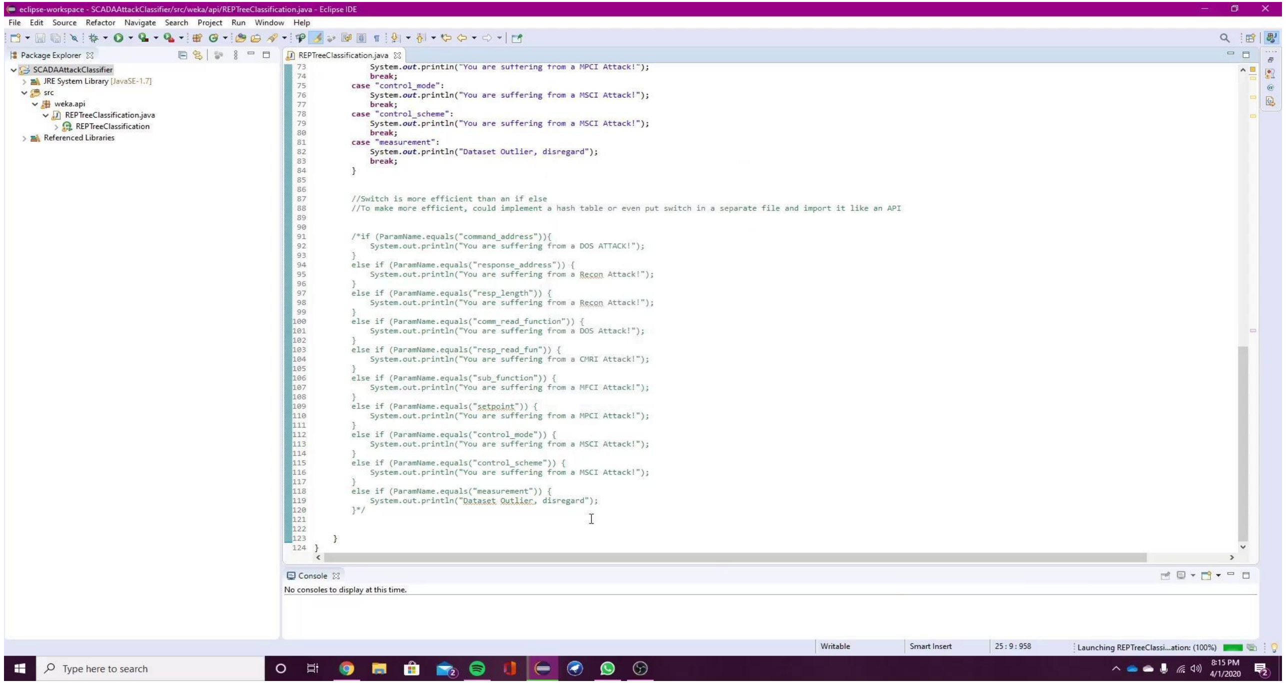
- ➔ We will be using a data processing tool, Weka, to classify our dataset.
- ➔ The machine learning algorithm being used to classify will be the REPTree to better make decisions of the type of attack.
- ➔ Once the ML algorithm determines the attack based on our data points, we move into the framework.
- ➔ The ATT&CK for ICS framework will assist in determining the type of attack tactic as well as the severity level.
- ➔ We will combine all of this to create an automated alerting system for SCADA and Industrial Control systems.

See Our Model in ACTION!

Potential Next Steps, (if endorsed)

- ➔ Create a high/mid-level prototype of the alert system including the following features:
  - ◆ User-friendly GUI
  - ◆ Immediate alert system to end-user
  - ◆ Sustainability across different OS and platforms
  - ◆ Adaptability to different industries

MVP (Minimum Viable Product)



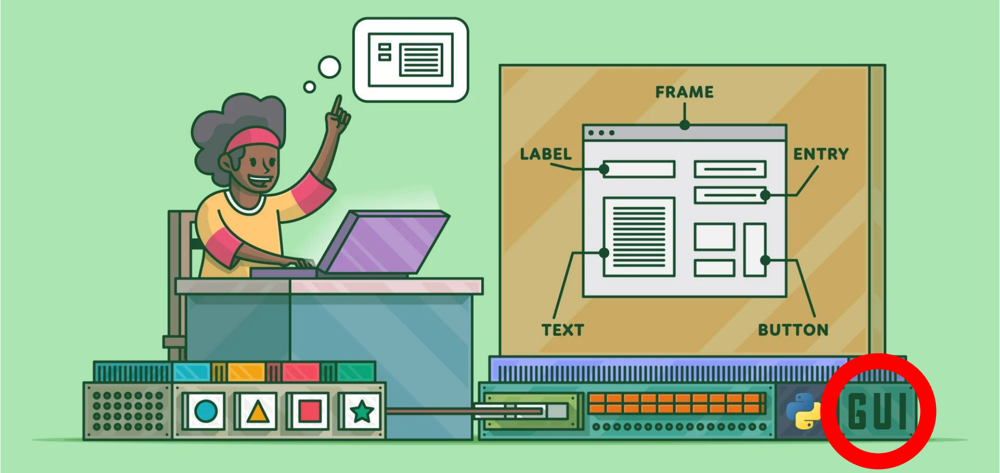
```java
            System.out.println("You are suffering from a MPCI Attack!");
            break;
        case "control_mode":
            System.out.println("You are suffering from a MSCI Attack!");
            break;
        case "control_scheme":
            System.out.println("You are suffering from a MSCI Attack!");
            break;
        case "measurement":
            System.out.println("Dataset Outlier, disregard");
            break;
        }


        //Switch is more efficient than an if else
        //To make more efficient, could implement a hash table or even put switch in a separate file and import it like an API


        /*if (ParamName.equals("command_address")){
            System.out.println("You are suffering from a DOS ATTACK!");
        }
        else if (ParamName.equals("response_address")) {
            System.out.println("You are suffering from a Recon Attack!");
        }
        else if (ParamName.equals("resp_length")) {
            System.out.println("You are suffering from a Recon Attack!");
        }
        else if (ParamName.equals("comm_read_function")) {
            System.out.println("You are suffering from a DOS Attack!");
        }
        else if (ParamName.equals("resp_read_fun")) {
            System.out.println("You are suffering from a CMRI Attack!");
        }
        else if (ParamName.equals("sub_function")) {
            System.out.println("You are suffering from a MFCI Attack!");
        }
        else if (ParamName.equals("setpoint")) {
            System.out.println("You are suffering from a MPCI Attack!");
        }
        else if (ParamName.equals("control_mode")) {
            System.out.println("You are suffering from a MSCI Attack!");
        }
        else if (ParamName.equals("control_scheme")) {
            System.out.println("You are suffering from a MSCI Attack!");
        }
        else if (ParamName.equals("measurement")) {
            System.out.println("Dataset Outlier, disregard");
        }*/


    }
}
```

No consoles to display at this time.

Development Options

Web-Based Application

Working with various Industries



Please direct any questions, concerns, or comments to pb10842p@pace.edu

We will appreciate any and all constructive feedback!

(Team 5 - Machine Learning for Information Assurance in SCADA Systems)