Industrial Power Control System Cyber Attacks Detection

November 30, 2022

Author:
Moriya Bitton
Victor Kushnir

1 Abstract

Industrial control systems (ICSs) are widely used and vital to industry and society. Their failure can have a severe impact on both economics and human life. Hence, these systems have become an attractive targets for attacks, both physical and cyber. Currently, the power system operators are heavily relied on to make decisions regarding the causes of experienced disturbances and the appropriate course of action as a response. In the case of cyber-attacks against a power system, human judgment is less certain since there is an overt attempt to disguise the attack and deceive the operators as to the true state of the system. To enable the human decision maker, we explore the viability of machine learning as a means for discriminating types of power system disturbances and focus specifically on detecting cyber-attacks where deception is a core idea of the event. In this project, we will try to build an ML model for the automatic detection of Industrial Power Control System Cyber Attacks.

2 Introduction

Industrial control systems (ICSs), also known as supervisory control and data acquisition (SCADA) systems, combine distributed computing with physical process monitoring and control. They are comprised of elements providing feedback from the physical world (sensors), elements influencing it (actuators), as well as computers and controller networks that process the feedback data and issue commands to the actuators. Many ICSs are safety-critical, and an attack interfering with their functionality can cause substantial financial and environmental harm, and endanger people's lives. The importance of ICSs makes them an attractive target for attacks, particularly cyber attacks. Several high impact incidents of this kind have been reported in recent years, including the attack on a power grid in Ukraine , the infamous Stuxnet malware that targeted nuclear centrifuges in Iran, and attacks on a Saudi oil company. In the past, ICSs ran on proprietary hardware and software in physically secure locations, but more recently they have adopted common information technology stack and remote connectivity. This trend exposes ICSs to cyber threats that leverage common technology stack attack tools. At the same time, the ICS defender's toolbox is limited due to the need to support legacy protocols built without modern security features, as well as the inadequate processing capabilities of the endpoints. This problem can be addressed by utilizing traditional IT network-based intrusion detection systems (IDS) to identify malicious activity, which does not rely on endpoint computational resources. However, the very low number of known attacks on ICSs renders this approach ineffective.

2.1 Synchrophasor based Smart Grid

The smart grid consists of two layers, cyber and physical systems. The two layers are coupled with each other and form the cyber-physical environment. The Synchrophasor or Phasor Measurement Unit (PMU) technology is built upon the cyber layer and provides real-time data to the energy management system (EMS) for the purpose of controlling the physical system. Such processes are presented as a sequence of execution events in the cyber-physical environment. The synchrophasor data includes not only the measurements such as voltage and current phasors but also the status of system devices including relays, breakers, switches, and transformers. The extreme low latency offered by time-synchronized data provides a huge volume of data with extra information and enables various real-time power system control algorithms in order to increase smart grid reliability and stability. The deployment of synchrophasor technology accelerates the use of communication networks within utilities and between neighboring utilities. The latest synchrophasor devices are vulnerable to cyber-attacks. there are still large numbers of legacy devices in service with little or no protection against the attacks.

Contemporary attacks against a power system can be launched from a compromised personal computer (PC) through a network to control a breaker. For example, the Aurora event highlights the potential for an attacker to open and close a breaker at high speed from a remote connection to damage an electric generator. Vulnerabilities can also be exploited against Intelligent Electronic Devices (IED) by uploading malicious settings. The Stuxnet worm is an example of settings changes on a control device causing a physical system to malfunction. Moreover, most network protocols used in power systems are open standard protocols without any security features. Such protocols include IEEE C37.118 protocol, used for synchrophasor data streaming, MODBUS, used to remotely monitor and control IED, and DNP3, which is also used to remotely monitor and control IED. The penetration tests conducted before have shown that cyber-attacks targeted against substation computers and devices can lead to Denial of Service (DoS) by making communication with a device impossible or causing devices to crash or reset and therefore prevent real time monitoring and controlling of the power system.

2.2 Wide Area Monitoring Systems

Originally, intrusion detection systems were introduced to IT systems to detect activities that violate security policies. Intrusion detection systems (IDS) can be misuse-based or anomaly-based. Misuse and signature-based IDS identify well-defined patterns of known attacks and ignore undefined attacks. An anomaly-based IDS must consider a system's normal behavior in order to detect anomalies. Therefore, any deviation from normal behavior will be considered an intrusion.

2.3 Specification-based Intrusion Detection System

A probabilistic network can provide a clear semantic structure for extracting knowledge relevant to a particular domain. As they are capable of showing dependencies and interdependencies between variables, they can be used for diagnosis, learning, explanation, and a variety of other inference-related tasks. Bayesian networks are widespread among probabilistic networks because they provide explicit graphical representations of cause-and-effect reasoning. It can also represent causality, depending on how it is interpreted. Bayesian networks can be used for developing attack graphs that are used for assessing network vulnerability. An attack graph represents the causal relation between two nodes where the compromise of one will lead to the compromise of the other.

3 ICS Power Systems

Figure 6.1 shows the power system framework configuration used in generating these scenarios. In the network diagram, we have several components. Firstly, G1 and G2 are power generators. R1 through R4 are Intelligent Electronic Devices (IEDs) that can switch the breakers on or off. These breakers are labeled BR1 through BR4. We also have two lines. Line One spans from breaker one (BR1) to breaker two (BR2) and Line Two spans from breaker three (BR3) to breaker four (BR4). Each IED controls a single breaker. BR1 is controlled by R1, BR2 is controlled by R2, and so on. The IEDs use a distance protection scheme that trips the breaker on detected faults whether actually valid or faked since they have no internal validation to detect the difference. Operators can also manually issue commands to the IEDs R1 through R4 to manually trip the breakers BR1 through BR4. Manual override is used when performing maintenance on the lines or other system components.

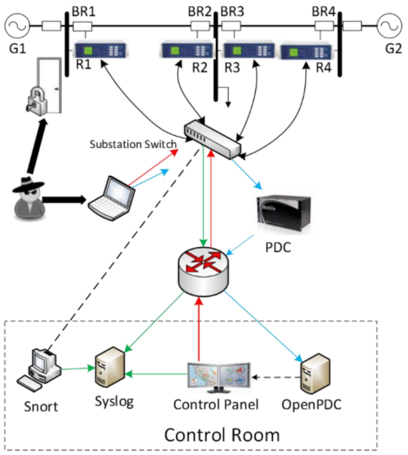


Figure 1: Power system framework configuration used in generating these scenarios

4 Types of Scenarios

1. Short-circuit fault – this is a short in a power line and can occur in various locations along the line, the location is indicated by the percentage range.

2. Line maintenance –one or more relays are disabled on a specific line to do maintenance for that line.

3. Remote tripping command injection (Attack) – this is an attack that sends a command to a relay which causes a breaker to open. It can only be done once an attacker has penetrated outside defenses.

4. Relay setting change (Attack) – relays are configured with a distance protection scheme and the attacker changes the setting to disable the relay function such that relay will not trip for a valid fault or a valid command.

5. Data Injection (Attack) – here we imitate a valid fault by changing values to parameters such as current, voltage, sequence components etc. This attack aims to blind the operator and causes a black out.

5 Related works

5.1 Power system datasets

The power system datasets have been used for multiple works related to power system cyber-attack classification.

Machine learning has distinguished itself as a discriminator of malicious and anomalous events in intrusion detection for traditional cyber security networks. These are systems that analyze the network transactions between computers and have been trained to characterize and recognize behavioral patterns in that traffic. Our approach is to extend this work and apply it to power systems, where networks are the means for communicating the state and operation of different power delivery components. This application focuses on the simultaneous assessment of dozens of variables associated with devices such as relays and generators as they are communicated within the power system network.

1. Pan, S., Morris, T., Adhikari, U., Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems, IEEE Transactions on Smart Grid. doi: 10.1109/TSG.2015.2409775.

2. Pan, S., Morris, T., Adhikari, U., Classification of Disturbances and Cyber-attacks in Power Systems Using Heterogeneous Time-synchronized Data, IEEE Transactions on Industrial Informatics. doi: 10.1109/TII.2015.2420951.

3. Pan, S., Morris, T., Adhikari, U., A Specification-based Intrusion Detection Framework for Cyber-physical Environment in Electric Power System, International Journal of Network Security (IJNS), Vol.17, No.2, PP.174-188, March 2015.

4. Beaver, J., Borges, R., Buckner, M., Morris, T., Adhikari, U., Pan, S., Machine Learning for Power System Disturbance and Cyber-attack Discrimination, Proceedings of the 7th International Symposium on Resilient Control Systems, August 19-21,2014, Denver, CO, USA.

People with different backgrounds have created various intrusion detection systems (IDS) that focus on different intrusions against Smart grid. One type of IDS research focuses on IED security within Smart grid. For example, Chee-Wooi Ten et al. has developed an anomaly-based detection technique for intrusions to IEDs. The Chee-Wooi Ten IDS is host-based and thus only identifies attacks against a single IED in the substation using sequential events recorded in the log from that IED. Another IDS proposed by Chen et al. Also, there is a known work that provides a protection mechanism for smart household appliances. Chen et al. created security rules for individual appliances by proposing homogeneous functions that models three factors of the appliance: device security, usability, and electricity pricing. More advanced IDS of this type will consider behaviors of multiple devices within the system to obtain system-level detection. Robert Mitchell et al. propose specification based IDS for the electric grid by considering the behaviors of three types of physical devices in the electric grid: head-ends, distribution access points/data aggregation points, and subscriber energy meters.

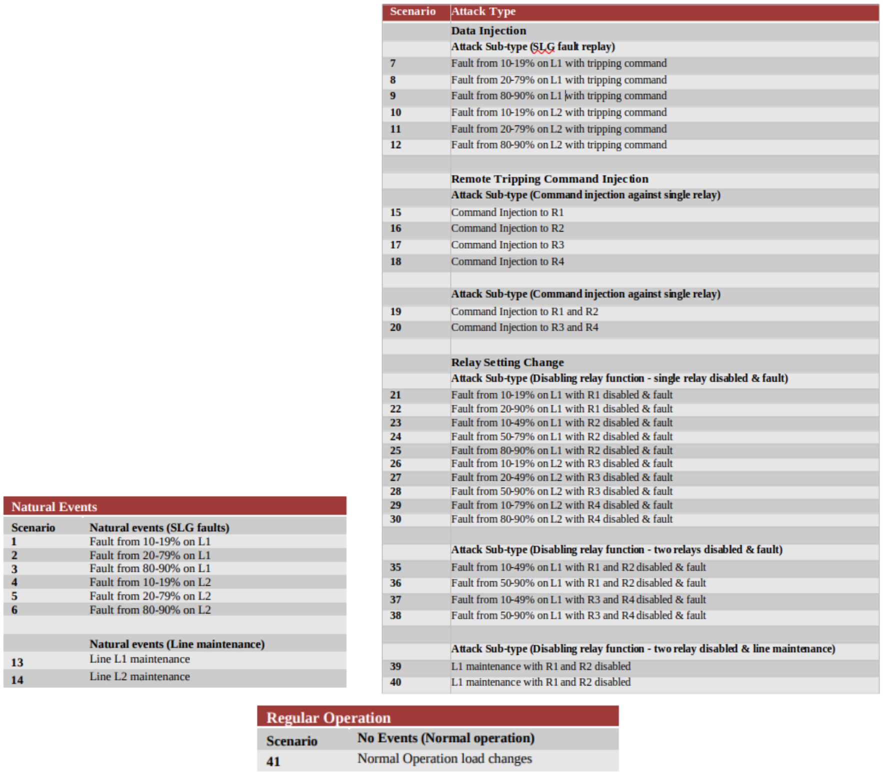
6 Data exploration

There are three datasets contained in this folder. They are made from one initial dataset consisting of fifteen sets with 37 power system event scenarios in each. The multi-class datasets are in ARFF format for easy use with Weka and the others are in CSV format also compatible with Weka. The 37 scenarios are divided into Natural Events (8), No Events (1) and Attack Events (28).
The datasets were randomly sampled at one percent and grouped into:

1. **Multi-class:**
   Figure 2 show the types of scenarios included.

2. **Three-class:**
   Figure 3 shows the distribution of instances in the three classification group.

3. **Binary:**
   Figure 4 shows the distribution of instances in the binary classification group.

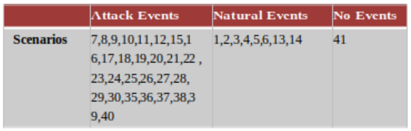


Scenario	Attack Type
	Data Injection
	Attack Sub-type (SLG fault replay)
7	Fault from 10-19% on L1 with tripping command
8	Fault from 20-79% on L1 with tripping command
9	Fault from 80-90% on L1 with tripping command
10	Fault from 10-19% on L2 with tripping command
11	Fault from 20-79% on L2 with tripping command
12	Fault from 80-90% on L2 with tripping command
	Remote Tripping Command Injection
	Attack Sub-type (Command injection against single relay)
15	Command Injection to R1
16	Command Injection to R2
17	Command Injection to R3
18	Command Injection to R4
	Attack Sub-type (Command injection against single relay)
19	Command Injection to R1 and R2
20	Command Injection to R3 and R4
	Relay Setting Change
	Attack Sub-type (Disabling relay function - single relay disabled & fault)
21	Fault from 10-19% on L1 with R1 disabled & fault
22	Fault from 20-90% on L1 with R1 disabled & fault
23	Fault from 10-49% on L1 with R2 disabled & fault
24	Fault from 50-79% on L1 with R2 disabled & fault
25	Fault from 80-90% on L1 with R2 disabled & fault
26	Fault from 10-19% on L2 with R3 disabled & fault
27	Fault from 20-49% on L2 with R3 disabled & fault
28	Fault from 50-90% on L2 with R3 disabled & fault
29	Fault from 10-79% on L2 with R4 disabled & fault
30	Fault from 80-90% on L2 with R4 disabled & fault
	Attack Sub-type (Disabling relay function - two relays disabled & fault)
35	Fault from 10-49% on L1 with R1 and R2 disabled & fault
36	Fault from 50-90% on L1 with R1 and R2 disabled & fault
37	Fault from 10-49% on L1 with R3 and R4 disabled & fault
38	Fault from 50-90% on L1 with R3 and R4 disabled & fault
	Attack Sub-type (Disabling relay function - two relay disabled & line maintenance)
39	L1 maintenance with R1 and R2 disabled
40	L1 maintenance with R1 and R2 disabled

Natural Events	
Scenario	**Natural events (SLG faults)**
1	Fault from 10-19% on L1
2	Fault from 20-79% on L1
3	Fault from 80-90% on L1
4	Fault from 10-19% on L2
5	Fault from 20-79% on L2
6	Fault from 80-90% on L2
	Natural events (Line maintenance)
13	Line L1 maintenance
14	Line L2 maintenance

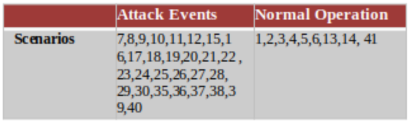
Regular Operation	
Scenario	**No Events (Normal operation)**
41	Normal Operation load changes

Figure 2: Multi-class



	Attack Events	Natural Events	No Events
Scenarios	7,8,9,10,11,12,15,1 6,17,18,19,20,21,22 , 23,24,25,26,27,28, 29,30,35,36,37,38,3 9,40	1,2,3,4,5,6,13,14	41

Figure 3: Three-class

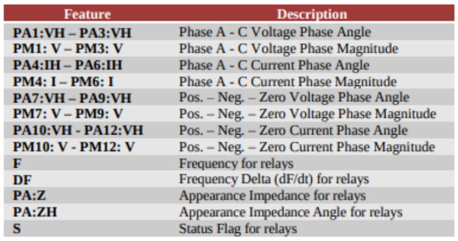


	Attack Events	Normal Operation
Scenarios	7,8,9,10,11,12,15,1 6,17,18,19,20,21,22 , 23,24,25,26,27,28, 29,30,35,36,37,38,3 9,40	1,2,3,4,5,6,13,14, 41

Figure 4: Binary classification

6.1 PMU

The 128 features are explained in the table below. There are 29 types of measurements from each phasor measurement units (PMU). A phasor measurement unit (PMU) or synchrophasor is a device which measures the electrical waves on an electricity grid, using a common time source for synchronization. In our system there are 4 PMUs which measure 29 features for 116 PMU measurement columns total. The index of each column is in the form of “R#-Signal Reference” that indicates a type of measurement from a PMU specified by “R#”. The signal references and corresponding descriptions are listed below. For example, R1-PA1:VH means Phase A voltage phase angle measured by PMU R1. After the PMU measurement columns, there are 12 columns for control panel logs, Snort alerts and relay logs of the 4 PMU/relay (relay and PMU are integrated together). The last column is the marker. The first three digits on the right is the load condition (in Megawatt). Another three digits to their left is fault locations, for example, “085” means fault at 85% of the transmission line specified by scenario description. However, for those that do not involve fault, e.g. “line maintenance”, these digits will be set to 000. The most left one digit or two digits indicate(s) the scenario number.



Feature	Description
PA1:VH – PA3:VH	Phase A - C Voltage Phase Angle
PM1: V – PM3: V	Phase A - C Voltage Phase Magnitude
PA4:IH – PA6:IH	Phase A - C Current Phase Angle
PM4: I – PM6: I	Phase A - C Current Phase Magnitude
PA7:VH – PA9:VH	Pos. – Neg. – Zero Voltage Phase Angle
PM7: V – PM9: V	Pos. – Neg. – Zero Voltage Phase Magnitude
PA10:VH - PA12:VH	Pos. – Neg. – Zero Current Phase Angle
PM10: V - PM12: V	Pos. – Neg. – Zero Current Phase Magnitude
F	Frequency for relays
DF	Frequency Delta (dF/dt) for relays
PA:Z	Appearance Impedance for relays
PA:ZH	Appearance Impedance Angle for relays
S	Status Flag for relays