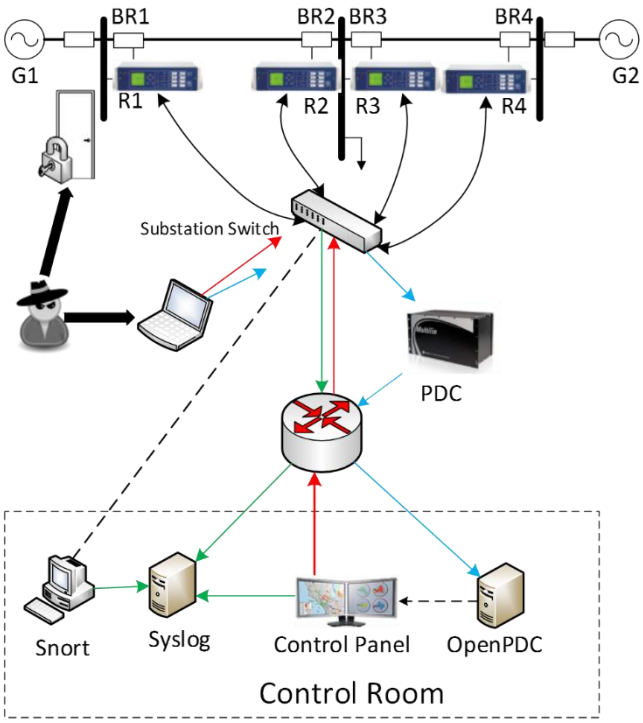
Power System Attack Datasets - Mississippi State University and Oak Ridge National Laboratory - 4/15/2014

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 multiclass 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:

- Binary
- Three-class and
- Multiclass datasets.

The figure below 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 automatically controls one breaker. R1 controls BR1, R2 controls BR2 and son on accordingly. The IEDs use a distance protection scheme which 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. The manual override is used when performing maintenance on the lines or other system components.



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.

- Tables I, II and III show the types of scenarios included.
- Table IV shows the distribution of instances in the binary classification group and
- Table V shows the distribution of instances in the binary classification group.

TABLE I. NATURAL EVENT SCENARIOS

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

TABLE II. NO EVENT SCENARIOS

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

TABLE III. ATTACK EVENT SCENARIOS

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

TABLE IV. THREE-CLASS CLASSIFICATION GROUP

	Attack Events	Natural Events	No Events
Scenarios	7,8,9,10,11,12,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,35,36,37,38,39,40	1,2,3,4,5,6,13,14	41

TABLE V. BINARY CLASSIFICATION

	Attack Events	Normal Operation
Scenarios	7,8,9,10,11,12,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,35,36,37,38,39,40	1,2,3,4,5,6,13,14, 41

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