HAI SECURITY DATASET

HIL-BASED AUGMENTED ICS (HAI) SECURITY DATASET WAS COLLECTED FROM A REALISTIC INDUSTRIAL CONTROL SYSTEM (ICS) TESTBED AUGMENTED WITH A HARDWARE-IN-THE-LOOP (HIL) SIMULATOR THAT EMULATES STEAM-TURBINE POWER GENERATION AND PUMPED-STORAGE HYDROPOWER GENERATION

Document version: 4.0
Release date: May 2023

RELEASE HISTORY

HAI is a security dataset that includes both the normal and abnormal behaviors for ICS anomaly detection research. The normal dataset was collected continuously for several days. Moreover, the abnormal dataset was collected based on various attack scenarios with the six feedback control loops in three different types of industrial control devices, namely the Emerson Ovation, GE Mark-VIe, and Siemens S7-1500. From the version 23.05, we also provide HAIEnd dataset that includes more detailed information about the internal control logic behaviors for Emerson boiler process control.

Version History

Four major versions of HAI dataset have been released until now. Each dataset consists of several CSV files, and each file satisfies time continuity. The quantitative summary of each version are as follows:

Release Version	# of tags	Normal Dataset			Abnormal Dataset			
		File (CSV)	Duration (hours)	Size (MB)	File (CSV)	# of attack	Duration (hours)	Size (MB)
HAIEnd 23.05 *HAI 23.05*	***225*** *86*	***end-train1*** *hai-train1*	*78*	***250.5*** *154.9*	***end-test1*** *hai-test1*	*14*	*15*	***48.2*** *29.8*
		end-train2 *hai-train2*	*81*	***260.7*** *161.3*	***end-test2*** *hai-test2*	*38*	*64*	***204.8*** *126.8*
		end-train3 *hai-train3*	*35*	***112.7*** *69.4*				
		end-train4 *hai-train4*	*55*	***176.0*** *109.2*				
		Sum	***249***	***799.9*** *494.8*	***Sum***	***52***	***79***	***253.0*** *156.6*
HAI 22.04	***86***	*train1*	*26*	*50.7*	*test1*	*7*	*24*	*48.2*
		train2	*56*	*108.9*	*test2*	*17*	*23*	*44.5*
		train3	*35*	*66.7*	*test3*	*10*	*17.3*	*33.4*
		train4	*24*	*45.7*	*test4*	*24*	*36*	*69.5*
		train5	*66*	*125.6*				-
		train6	*72*	*136.8*				
		Sum	***279***	***534.4***	***Sum***	***58***	***100.3***	***195.6***
HAI 21.03	***78***	*train1*	*60*	*110*	*test1*	*5*	*12*	*22*
		train2	*63*	*116*	*test2*	*20*	*33*	*61*
		train3	*229*	*245*	*test3*	*8*	*30*	*55*
					test4	*5*	*11*	*20*
					test5	*12*	*26*	*47*
		Sum	***352***	***471***	***Sum***	***50***	***112***	***205***
HAI 20.07	***59***	*train1*	*86*	*127*	*test1*	*28*	*81*	*119*
		train2	*91*	*98*	*test2*	*10*	*42*	*62*
		Sum	***177***	***225***	***Sum***	***38***	***123***	***181***

Note: 1) The version numbering follows a date-based scheme, where the version number indicates the released date of a HAI dataset. 2) HAI 23.05 has the same experimental configuration as that of 22.04, 3) Both HAI 23.05 and HAIEnd 23.05 data were collected simultaneously in the same experiment.

Document Change Logs

Version	Release Date	Changes	Page(s)
v4.0	May 31, 2023	***Major revision for HAI 23.05***	
		+ History and quantitative summary for HAI/HAIEnd 23.05	*01*
		+ Brief description of the boiler process control	*06*
		+ Detailed description of the control logic for pressure control	*08-09*
		+ 255 more data points for HAIEnd 23.05	*15-23*
		+ 16 more attack scenarios	*25-2*
		+ Details of HAI/HAIEnd 23.05	*30-32*
		+ Case studies with NetworkX graphs	*42-46*
		+ Citing datasets	*47*
v3.0	Apr. 29, 2022	***Major revision for HAI 22.04***	
		+ Version history for HAI 22.04	*01*
		+ Brief description of the boiler cooling system	*03-05*
		+ Detailed description of the boiler cooling controller	*08*
		+ 8 more data points	*12 – 14*
		+ 12 more attack scenarios	*25 – 28*
		+ Correct some errors on the attack scenarios	
		+ Details of HAI 22.04	*33 – 35*
		+ Citing datasets	*47*
v2.0	Feb. 17, 2021	***Major revision for HAI 21.03***	
		+ Brief description of the turbine trip control	*10*
		+ 19 more data points	*12 – 14*
		+ 11 more attack scenarios	*25 – 28*
		- Description related to multiple attacks	*25 – 28*
		+ Details of HAI 21.03	*36 – 38*
		+ Changes to HAI 20.07	*38 – 41*
v1.1	Jul. 22, 2020	***Minor revision for HAI 20.07***	
		+ New version numbering scheme	*All*
		+ Value ranges and description of data points	*12 – 15*
		+ Time duration in attack timetable	*39 – 40*
v1.0	Feb. 17, 2020	***Initial release for HAI v1.0 (20.02)***	*All*

HAI SECURITY DATASET

HIL-based augmented ICS (HAI) security dataset was collected from a realistic industrial control system (ICS) testbed augmented with a hardwire-in-the-loop (HIL) simulator that emulates steam-turbine power generation and pumped-storage hydropower generation.

Background

This dataset was developed for research on anomaly detection in cyber–physical systems (CPSs) such as railways, water treatment plants, and power plants.

In 2017, three laboratory-scale CPS testbeds were initially launched, namely GE's turbine testbed, Emerson's boiler testbed, and FESTO's modular production system (MPS) water treatment testbed. These testbeds were related to relatively simple processes, and were operated independent to each other. In September 2018, a complex process system was built to combine the three testbeds using a HIL simulator, where thermal power generation and pumped-storage hydropower generation were simulated. This ensured that the variables were highly coupled and correlated for a richer dataset. In addition, an open platform communications united architecture (OPC-UA) gateway was installed to facilitate data collection from heterogeneous devices.

The first version of the HAI dataset was made available on GitHub and Kaggle in February 2020. This dataset included ICS operational data from normal and abnormal situations for 38 attacks. Subsequently, a debugged version of HAI v1.0, namely HAI 20.07, was released in July 2020. We newly made HAI v2.0 for the HAICon 2020 competition and a refined version, namely HAI 21.03, was released in March 2021. In 2021, we held an AI-based competition named HAICon 2021. It was an AI-based challenge for industrial control system threat detection. We released the HAI 22.04 version refining the dataset used in the competition. In 2022, HAI and HAIEnd 23.05 were developed for ICS Endpoint threat detection.

HAI Testbed

The testbed consisted of a boiler, turbine, water-treatment component, and an HIL simulator. The boiler process involved water-to-water heat transfer based on low pressure and moderate temperature. On the other hand, the turbine process involved closely simulating the behavior of an actual rotating machine using a rotor kit testbed. The boiler and turbine processes were interconnected with the HIL simulator to ensure synchronization with the rotating speed of a steam-power generator. In the water treatment process, water was pumped to the upper reservoir and subsequently released into the lower reservoir according to a pumped-storage hydropower generation model during the HIL simulation.

The three real-world processes, that is, the boiler, turbine, and water treatment processes, were controlled by three different controllers. Emerson Ovation distributed control system (DCS) was used for controlling the water level, flow rate, pressure, temperature, water feed pump, and heater in the boiler process. In the turbine process, GE's Mark VIe DCS was used for speed control and vibration monitoring. A Siemens S7-300 PLC was used in the water treatment process to control the water level and pump. A dSPACE® SCALEXIO system was used for the HIL simulations and interconnected with the real-world processes using a Siemens S7-1500 PLC and ET200 remote IO devices.

HAI TESTBED

Process Architecture

The process flow of the testbed was divided into four primary processes: the boiler process (P1), turbine process (P2), water treatment process (P3), and HIL simulation (P4) (Figure 1). The HIL simulation enhances the correlation between the three real-world processes at the signal level by simulating the thermal power and pumped-storage hydropower generation processes.

The boiler and turbine processes simulated the thermal power plant, while the water treatment process simulated the pumped-storage hydropower plant.

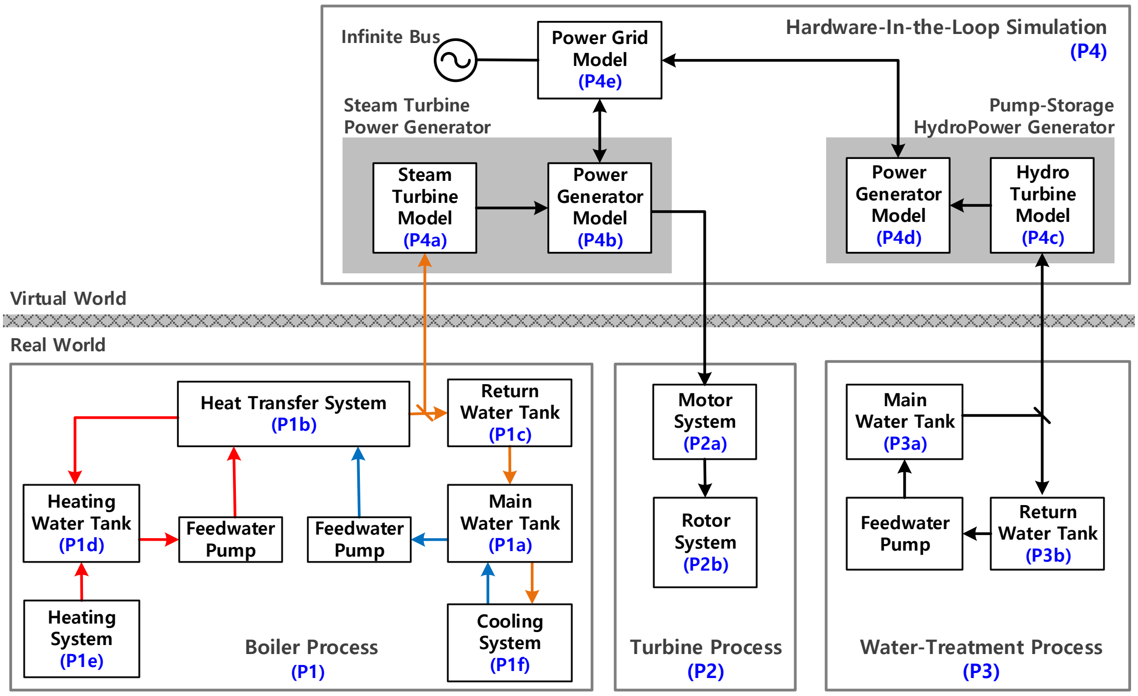

FIGURE 1. PROCESS FLOW DIAGRAM OF THE TESTBED.

P1: BOILER PROCESS

The boiler process involved water-to-water heat transfer at low pressures and moderate temperatures, where the boiler pressure, temperature, and water level are controlled by the boiler process. The opening and closing rates of the main valve are also controlled according to the opening rate of the steam valve of the thermal power plant in the HIL simulator. The pressure and temperature of the main pipe and the water level are transmitted to the HIL simulator in real-time to determine the amount of power generated.

Cool water in the main water tank (P1a) is pumped to the heat-transfer system (P1b) through a feedwater pump, subsequently providing water at a constant temperature and pressure to the return water tank (P1c). The heating system (P1e) transfers thermal energy through the water to the heat transfer system. The water temperature and pressure values are then converted into the current steam temperature and pressure values for the steam-turbine power generator of the HIL simulator (P4a). Water flows from the return water tank (P1c) to the main water tank (P1a) at a constant flow rate, thereby maintaining constant water level in the return water tank. The water circulating to the main tank is not sufficiently cooled; therefore, the cooling system (P1f) additionally removes the thermal energy from the water in the main water tank. The temperature, pressure, level, and flow rate of water in the boiler system were kept constant using eleven sensors, three actuators (two pumps and a heater), and six valves. An operator was able to control five setpoints via the operator workstation (OWS). Boiler process

P2: TURBINE PROCESS

An actual rotating machine was closely simulated using a GE Rotor Kit (Bently Nevada Asset Condition Monitoring), which consisted of a motor system with a direct-current motor speed control device and a rotor system that allows for coupling and included a rotor shaft, two balance wheels, two journal bearings, and a bearing block. The motor speed was synchronized with the rotating speed of the thermal power generator model in the HIL simulator. The turbine system included a speedometer and four vibration-monitoring proximity probes to maintain a motor speed constant, where the operator can adjust the turbine rotations per minute (RPM) setpoint using a human-machine interface (HMI).

P3: WATER-TREATMENT PROCESS

The water-treatment process involved the pumping and release of water between the upper and lower reservoirs using the hydropower turbine model in the HIL simulation. The water-treatment system included seven sensors, one actuator, and an outflow control valve to control the flow and pressure from the return water tank (P3b) to the main water tank (P3a), as well as the water level in the main water tank. The hydraulic pressure, flow rate, and water level of the upper water tank were transmitted to the HIL simulator in real time to determine the power generation.

P4: HARDWARE-IN-THE-LOOP SIMULATOR

The simulation system consisted of two synchronous generator models (*i.e., steam-turbine power generator and pumped-storage hydropower generator*) and one power grid model, which included the local load demand and was connected to an infinite bus.

An HIL-based simulator was developed to combine the three control systems for the boiler, turbine, and water treatment processes to form a combined power generation system. Specifically, the temperature and pressure of the boiler system were used to determine the pressure and temperature of the steam entering the steam turbine model (STM) (P4.1). The output power of the STM was controlled by an internal steam governor, and the power generator model (P4.2) generated the corresponding electrical power. Further, the hydro turbine model (HTM) (P4.3) and power generator model (P4.4) calculated the generated output power based on the discharge from the water treatment system, where both models were controlled to ensure that the frequency of the microgrid load was 60 Hz (P4.5). The power generated based on the input load was dependent on the opening and closing rates of the valves of the thermal power plant and pumped-storage power plant. Thus, the opening and closing rates of the valves in the control systems for the boiler and water treatment systems were determined.

Testbed Components

The three real-world processes were controlled by three different controllers. Specifically, the boiler process was controlled by Emerson's Ovation DCS for the water level, flow rate, pressure, temperature, water feed pump, and heater control. The turbine process was controlled by GE's Mark VIe DCS for speed control and vibration monitoring, and the water treatment process was controlled by a Siemens S7-300 PLC for water level and pump control. In the HAI testbed, the HIL simulations were conducted using a dSPACE® SCALEXIO system interconnected with the real-world processes using a S7-1500 PLC (Siemens) and with an ET200 remote IO devices.

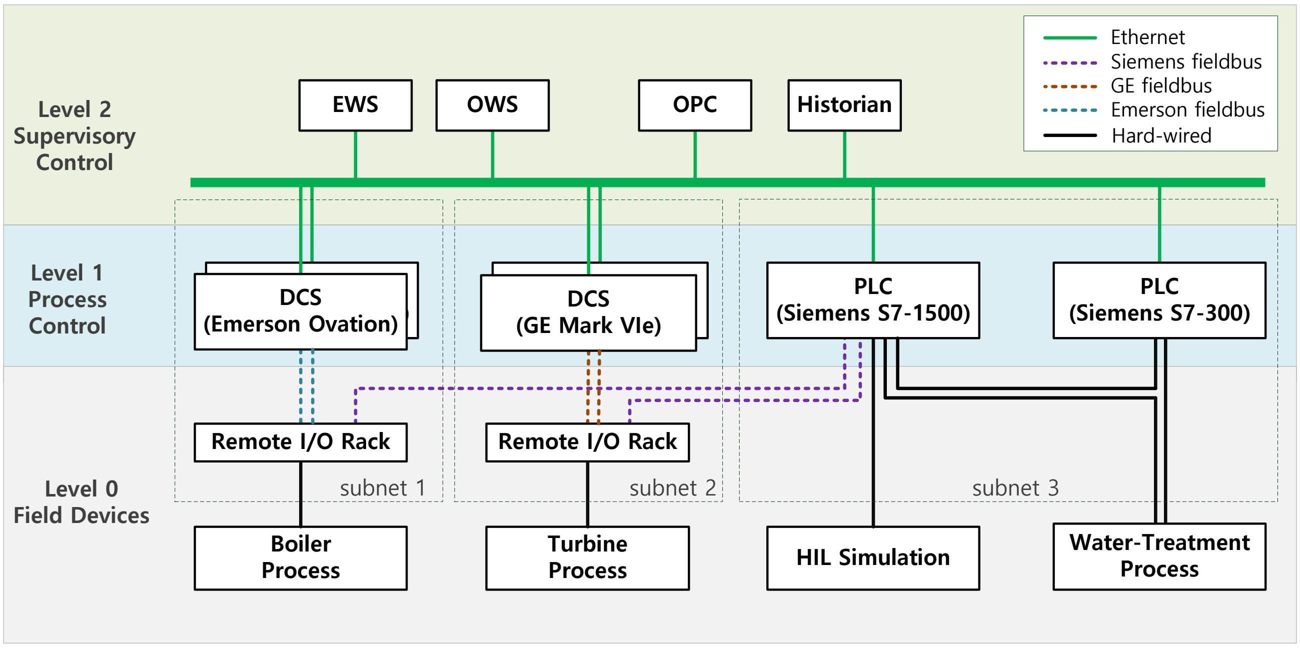

FIGURE 2. TESTBED COMPONENTS AND DATA FLOW.

BOILER PROCESS CONTROL

Emerson Ovation DCS consists of four feedback control loops to control the pressure, water level, outflow, temperature, and cooling pump.

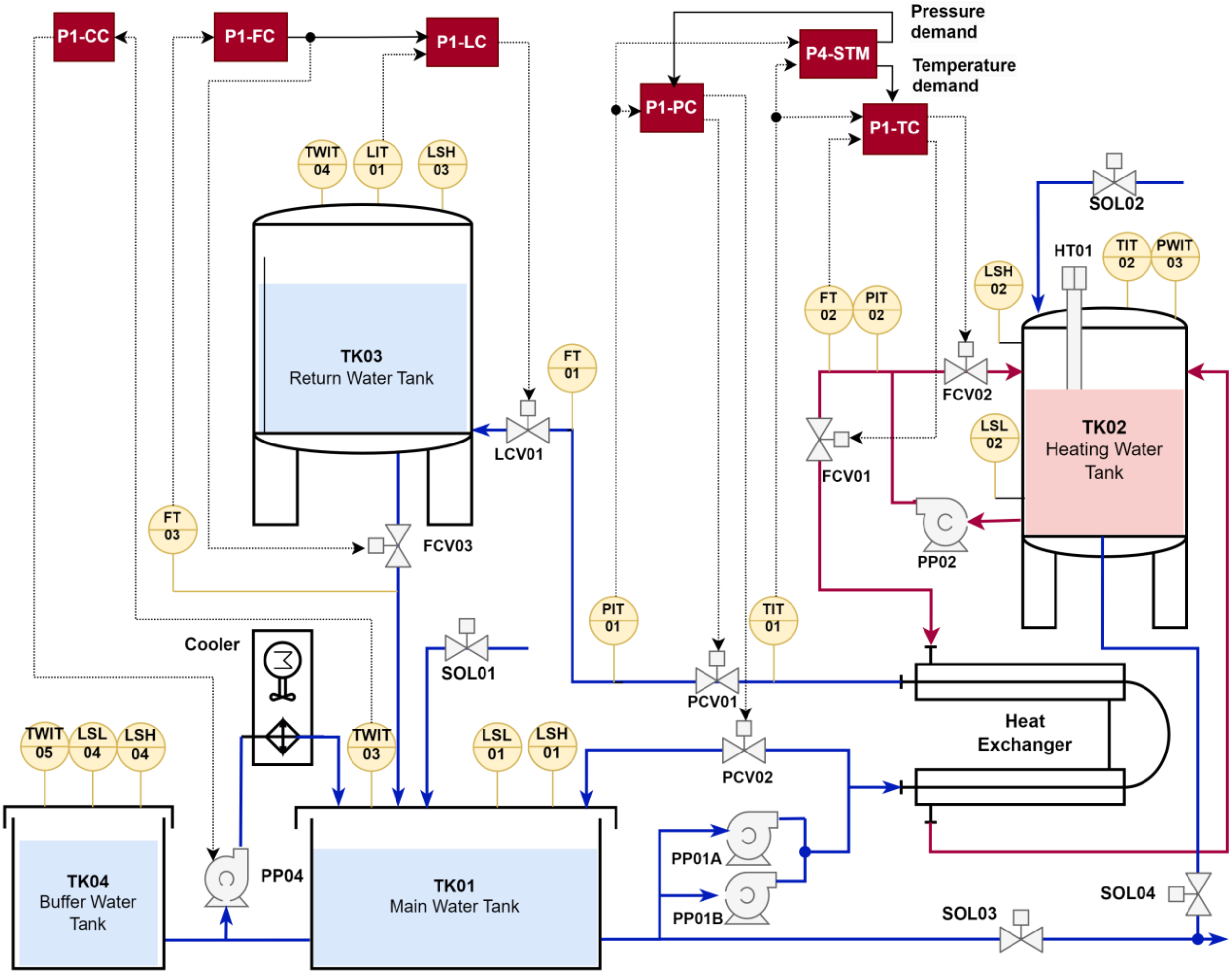

FIGURE 3. BOILER PROCESS CONTROL ARCHITECTURE.

P1-PC: Pressure Control

P1-PC is a feedback control loop for two pressure-control valves (PCV01D and PCV02D) and maintain the pressure (PIT01) between the main and return water tanks according to an operator's setpoint command (B2016).

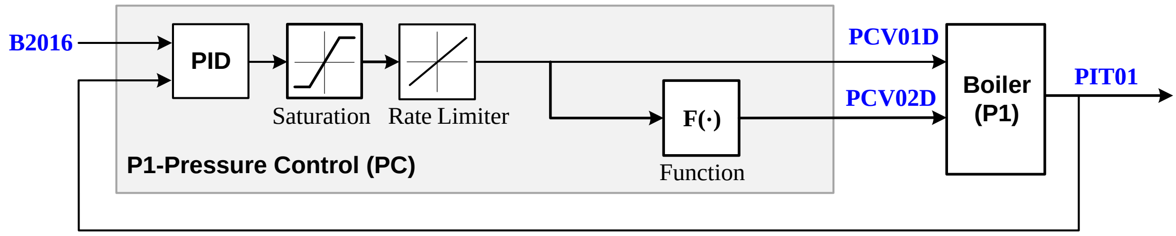

FIGURE 4. PRESSURE CONTROL OF THE BOILER.

P1-LC: Level Control

P1-LC is a feedback control loop for the level-control valve (LCV01D) and maintain the water level (LIT01) of the return water tank according to the operator's setpoint command (B3004). In addition, a feed-forward control was used to rapidly suppress any disturbance in the outflow rate (FCV03D).

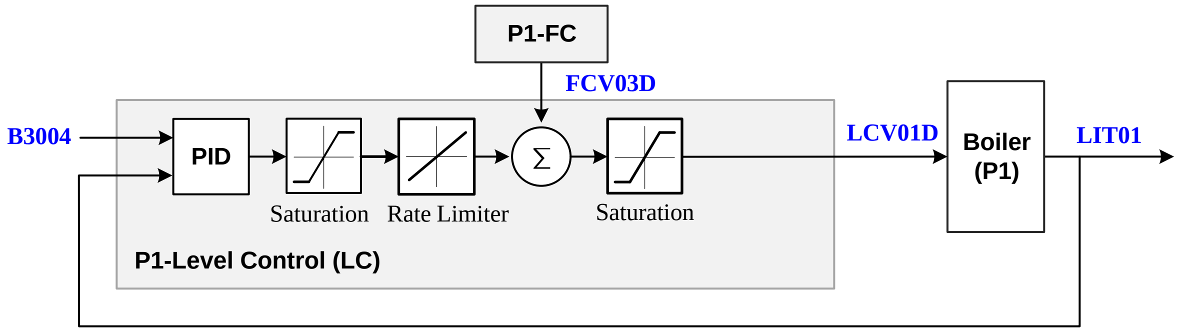

FIGURE 5. LEVEL CONTROL OF THE BOILER.

P1-FC: Flow rate Control

P1-FC is a feedback control loop for the flow-control valve (FCV03D) and maintain the outflow rate (FT03) for the return water tank according to the operator's setpoint command (B3005).

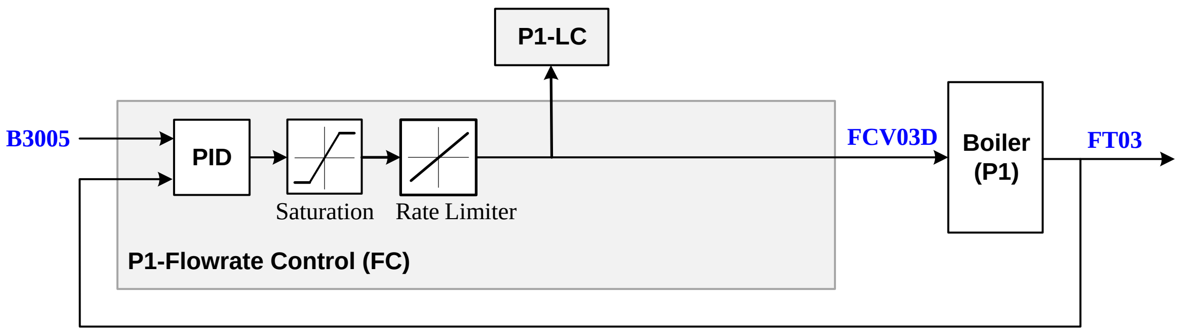

FIGURE 6. FLOW RATE CONTROL OF THE BOILER.

P1-TC: Temperature Control

P1-TC is a feedback control loop for two flow-control valves (FCV01D and FCV02D) in the heat transfer system and maintain the temperature (TIT01) of the main vessel according to the operator's setpoint command (B4022). Cascade control with feedforward compensation to the flow controller (inner loop) based on the water flow allowed for a quicker response to fluctuations in the water flow.

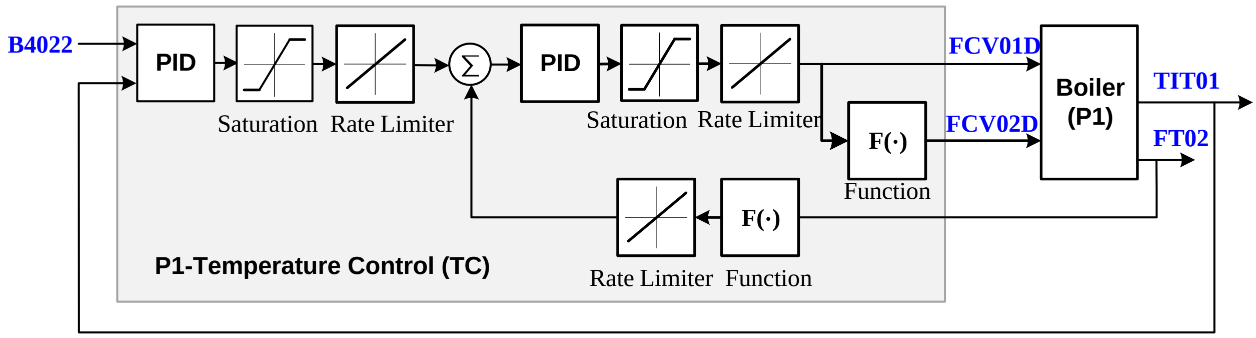

FIGURE 7. TEMPERATURE CONTROL OF THE BOILER.

P1-CC: Cooling Control

P1-CC drives frequency (PP04) of the cooling water pump. This activates the pump operation at the set point (PP04SP) when the water temperature (TIT03) in the main water tank is in the operation range.

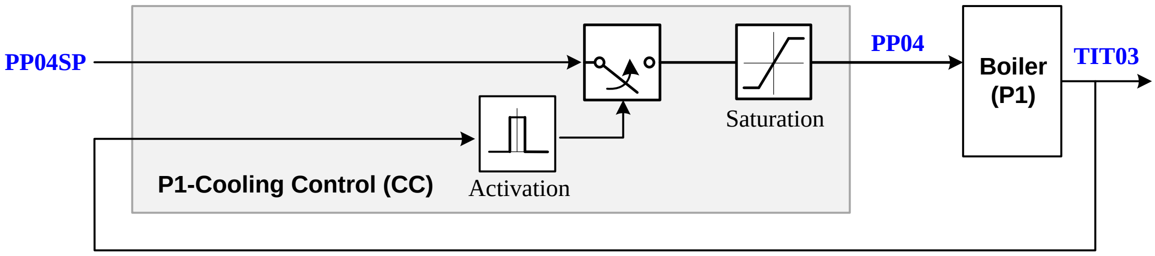

FIGURE 8. COOLING CONTROL OF THE BOILER

BOILER CONTROL LOGICS

The HAIEnd dataset covered control loop of Emerson Ovation DCS in detail. For the sake of better understanding, we additionally provide the detailed control logic of boiler process.

P1-PC Control Logics

P1-PC is a feedback control loop for two pressure-control valves (PCV01D and PCV02D) and maintain the pressure (PIT01) between the main and return water tanks according to an operator's setpoint command (B2016). HAI dataset included only I/O of control process.

In fact, the control loop is not a single logic but a collection of multiple algorithm blocks that perform different functions, such as a fast Boolean, flip-flop, and PID. For example, the I/O and internal points can be represented simultaneously by expressing algorithm functions as individual nodes in the control logic as shown in Figure 9. The control logic with I/O and internal points as nodes takes the form of a bidirectional graph. Each point name is marked on the edge. HAIEnd dataset contains all the named edges on the graph, as well as some point not connected to the control logic for maintenance. For more information, please refer to graph configuration files included in HAIEnd dataset.

HAIEnd dataset included both I/O (PCV01D, PCV02D, PIT01, B2016, PP01A, PP01B) and internal point as represented by the edge in control logic. Internal point is used to deliver processed value to each algorithm blocks.

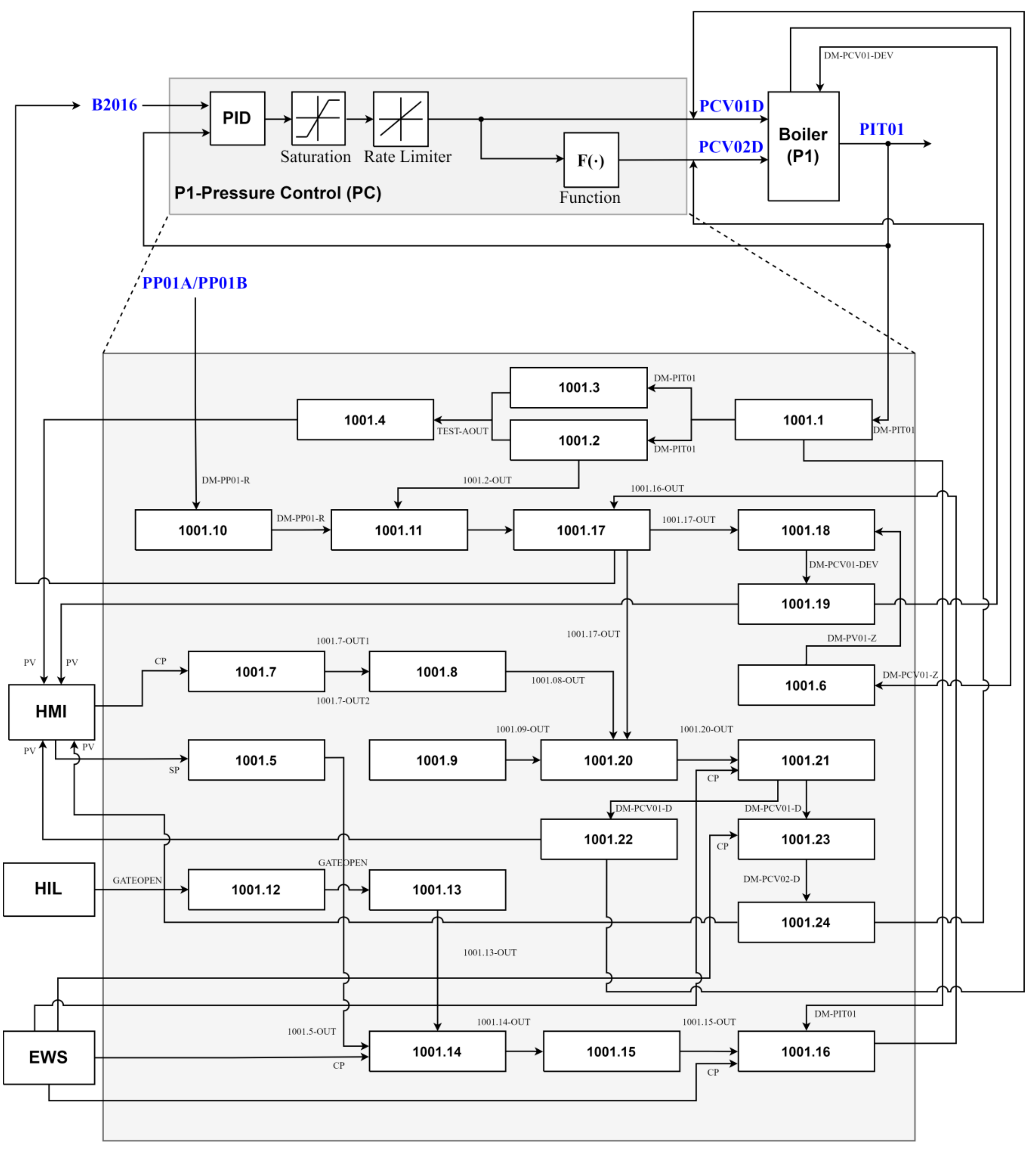

FIGURE 9. CONTROL LOGIC FOR PRESSURE CONTROL OF THE BOILER

TURBINE PROCESS CONTROL

GE's Mark VIe DCS has one feedback loop that controlled the motor speed. The HIL simulator (P4-STM) generate setpoint trajectories for speed control (P2-SC).

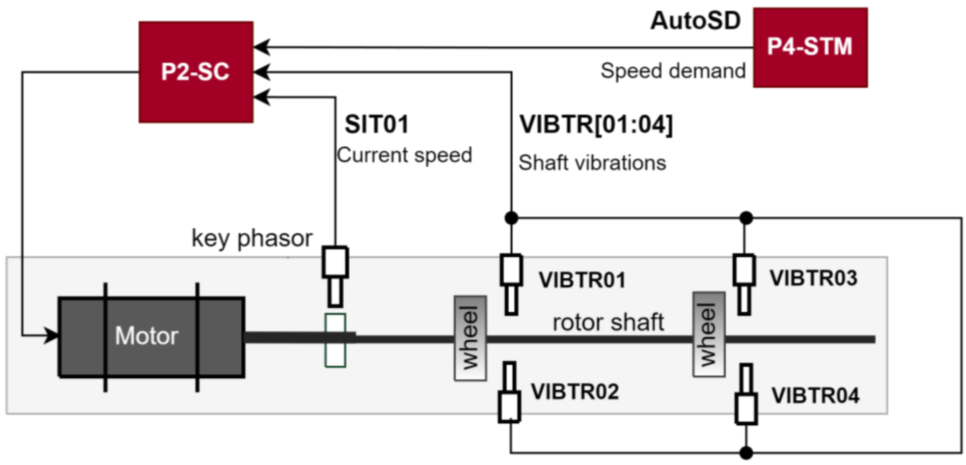

FIGURE 10. TURBINE PROCESS CONTROL ARCHITECTURE.

P2-TRIP: Over-speed and over-vibration trips

The purpose of trip is to prevent an over-speed and over-vibration of a turbine. A turbine runs when the monitored speed (SIT01) is above the RPM trip rate (RTR) or any of four vibration sensors (VIBTR[n]) are above a preset limit (VTR[n]), and then the emergency stop (Emerg) become active. The turbine run mode is activated if the push button to exit the trip mode (TripEx) is successfully triggered.

P2-SC: Speed Control

The P2-SC speed controller increases the motor speed from zero to the minimum controlling speed at a constant rate. Moreover, it facilitates engagement control with a proportional integral derivative (PID) controller to maintain the motor speed value (SIT01) as close as possible to the speed setpoint value (AutoSD).

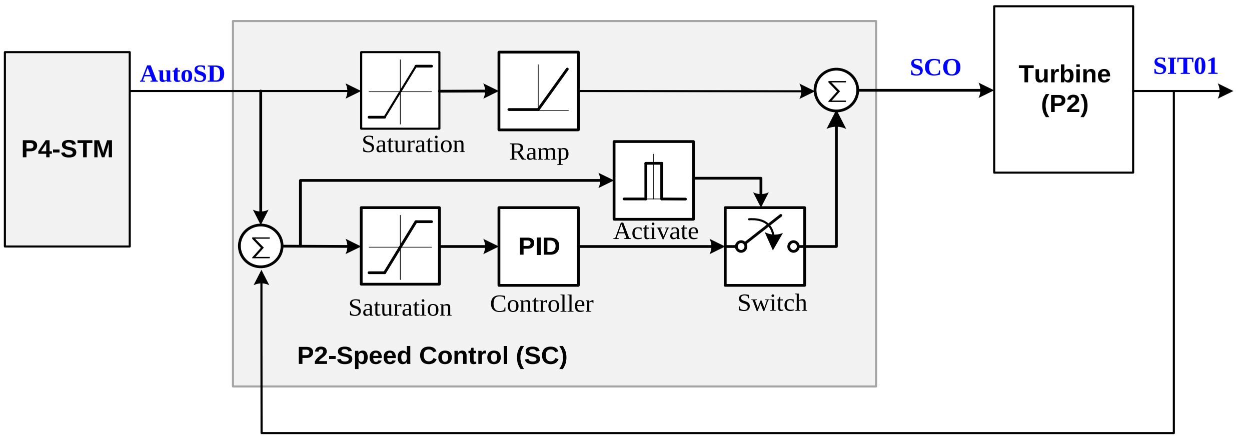

FIGURE 11. SPEED CONTROL OF A TURBINE.

WATER TREATMENT CONTROL

The SIMATIC S7 PCL used for the water treatment control has one feedback loop that controls the water level in the upper reservoir.

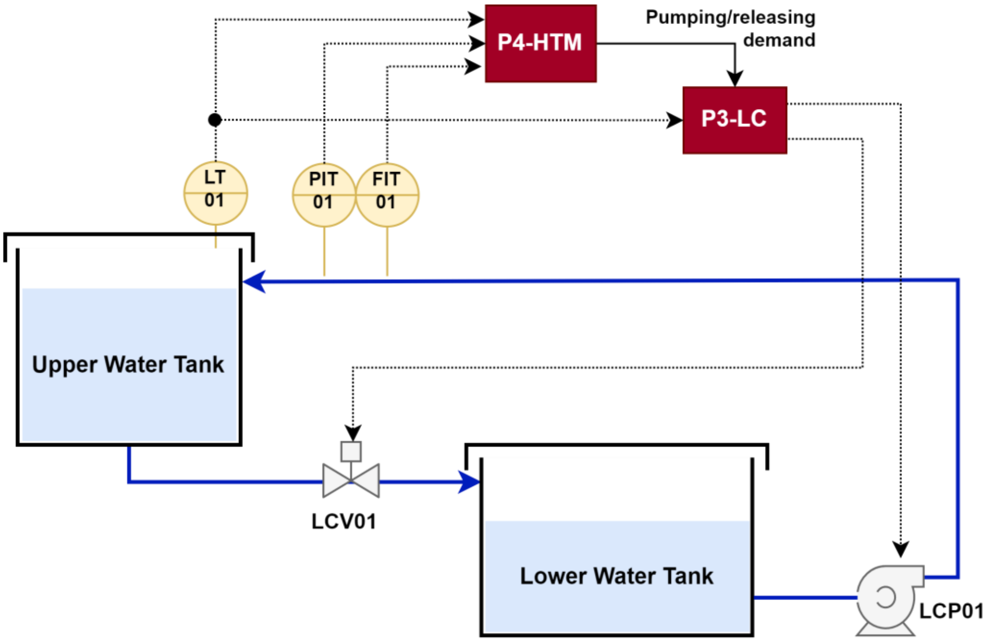

FIGURE 12. WATER TREATMENT PROCESS CONTROL ARCHITECTURE.

P3-LC: Level Control

P3-LC controls the level control valve (LCV01) and level control pump (LCP01) by adjusting the discharge and pumping demands of the HIL simulator.

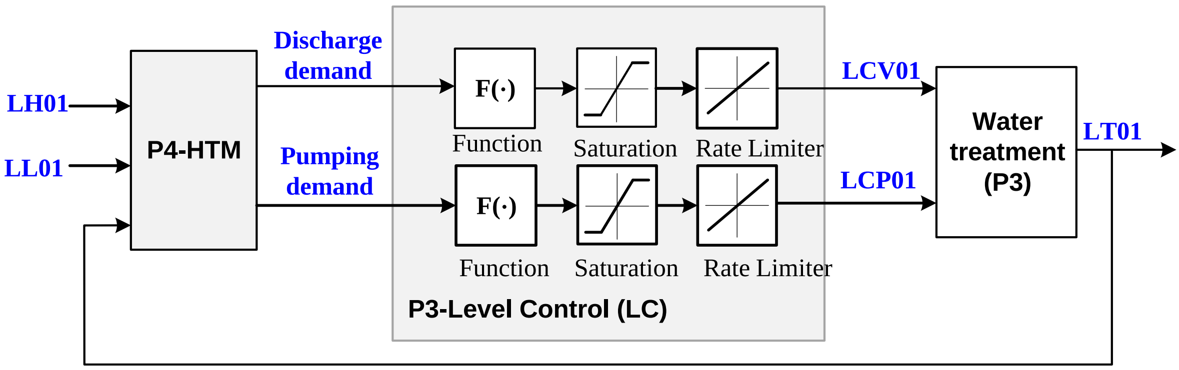

FIGURE 13. WATER LEVEL CONTROL IN A WATER TREATMENT PLANT.

Data Points

Supervisory control and data acquisition (SCADA) system typically consist of data elements called points (or tags), where each point represents a single variable measured or controlled by the system. The HAI dataset includes the critical data points to control and monitor at a centralized place. The HAIEnd dataset, however, internal points used in DCS logics to control the boiler process.

HAI DATA POINTS

As the HAI version becomes more recent, the number of data points are increases from 59 to 86. All data points of each version are tabulated below.

No	Name	Range		Unit	Description	HAI		
		Min	Max			20.07	21.03	22.04 23.05
1	P1_B2004	0	10	bar	Heat-exchanger outlet pressure setpoint	√	√	√
2	P1_B2016	0	10	bar	Pressure demand for thermal power output control	√	√	√
3	P1_B3004	0	720	mm	Water level setpoint (return water tank)	√	√	√
4	P1_B3005	0	2,500	l/h	Discharge flowrate setpoint (return water tank)	√	√	√
5	P1_B4002	0	100	℃	Heat-exchanger outlet temperature setpoint	√	√	√
6	P1_B4005	0	100	%	Temperature PID control output	√	√	√
7	P1_B400B	0	2,500	l/h	Water outflow rate setpoint (heating water tank)	√	√	√
8	P1_B4022	0	40	℃	Temperature demand for thermal power output control	√	√	√
9	P1_FCV01D	0	100	%	Position command for the FCV01 valve	√	√	√
10	P1_FCV01Z	0	100	%	Current position of the FCV01 valve	√	√	√
11	P1_FCV02D	0	100	%	Position command for the FCV02 valve	√	√	√
12	P1_FCV02Z	0	100	%	Current position of the FCV02 valve	√	√	√
13	P1_FCV03D	0	100	%	Position command for the FCV03 valve	√	√	√
14	P1_FCV03Z	0	100	%	Current position of the FCV03 valve	√	√	√
15	P1_FT01	0	2,500	mmH2O	Measured flowrate of the return water tank	√	√	√
16	P1_FT01Z	0	3,190	l/h	Water inflow rate converted from P1_FT01	√	√	√
17	P1_FT02	0	2,500	mmH2O	Measured flowrate of heating water tank	√	√	√
18	P1_FT02Z	0	3,190	l/h	Water outflow rate conversion from P1_FT02	√	√	√
19	P1_FT03	0	2,500	mmH2O	Measured flowrate of the return water tank	√	√	√
20	P1_FT03Z	0	3,190	l/h	Water outflow rate converted from P1_FT03	√	√	√
21	P1_LCV01D	0	100	%	Position command for the LCV01 valve	√	√	√
22	P1_LCV01Z	0	100	%	Current position of the LCV01 valve	√	√	√

No	Name	Range		Unit	Description	HAI		
		Min	Max			20.07	21.03	22.04 23.05
23	P1_LIT01	0	720	mm	Water level of the return water tank	√	√	√
24	P1_PCV01D	0	100	%	Position command for the PCV01 valve	√	√	√
25	P1_PCV01Z	0	100	%	Current position of the PCV01 valve	√	√	√
26	P1_PCV02D	0	100	%	Position command for the PCV2 valve	√	√	√
27	P1_PCV02Z	0	100	%	Current position of the PCV02 valve	√	√	√
28	P1_PIT01	0	10	bar	Heat-exchanger outlet pressure	√	√	√
29	P1_PIT01_HH	0	10	bar	Highest outlet pressure of the heat-exchanger			√
30	P1_PIT02	0	10	bar	Water supply pressure of the heating water pump	√	√	√
31	P1_PP01AD	0	1	Boolean	Start command of the main water pump PP01A		√	√
32	P1_PP01AR	0	1	Boolean	Running state of the main water pump PP01A		√	√
33	P1_PP01BD	0	1	Boolean	Start command of the main water pump PP01B		√	√
34	P1_PP01BR	0	1	Boolean	Running state of the main water pump PP01B		√	√
35	P1_PP02D	0	1	Boolean	Start command of the heating water pump PP02		√	√
36	P1_PP02R	0	1	Boolean	Running state of the heating water pump PP02		√	√
37	P1_PP04	0	100	%	Control out of the cooler pump			√
38	P1_PP04SP	0	100	℃	Cooler temperature setpoint			√
39	P1_SOL01D	0	1	Boolean	Open command of the main water tank supply valve			√
40	P1_SOL03D	0	1	Boolean	Open command of the main water tank drain valve			√
41	P1_STSP	0	1	Boolean	Start/stop command of the boiler DCS		√	√
42	P1_TIT01	-50	150	℃	Heat-exchanger outlet temperature	√	√	√
43	P1_TIT02	-50	150	℃	Temperature of the heating water tank	√	√	√
44	P1_TIT03	-50	150	℃	Temperature of the main water tank			√
45	P2_24Vdc	0	30	Voltage	DCS 24V Input Voltage	√	√	√
46	P2_ATSW_Lamp	0	1	Boolean	Lamp of the Auto SW			√
47	P2_AutoGo	0	1	Boolean	Auto start button	√ (Auto)	√	√
48	P2_AutoSD	0	3,200	RPM	Auto speed demand	√ (SD01)	√	√
49	P2_Emerg	0	1	Boolean	Emergency button	√ (Emgy)	√	√
50	P2_MASW	0	1	Boolean	Manual(1)/Auto(0) SW			√

No	Name	Range		Unit	Description	HAI		
		Min	Max			20.07	21.03	22.04 23.05
51	P2_MASW_Lamp	0	1	Boolean	Lamp of Manual SW			√
52	P2_ManualGO	0	1	Boolean	Manual start button		√	√
53	P2_ManualSD	0	3,200	RPM	Manual speed demand		√	√
54	P2_OnOff	0	1	Boolean	On/off switch of the turbine DCS	√ (On)	√	√
55	P2_RTR	0	2,880	RPM	RPM trip rate		√	√
56	P2_SCO	0	100,000	-	Control output value of the speed controller		√	√
57	P2_SCST	-100	100	RPM	Speed change proportional to frequency change of the STM		√	√
58	P2_SIT01	0	3,200	RPM	Current turbine RPM measured by speed probe	√	√	√
59	P2_TripEx	0	1	Boolean	Trip emergency exit button	√	√	√
60	P2_VIBTR01	-10	10	㎛	Shaft-vibration-related Y-axis displacement near the 1st mass wheel	√ (VYT02)	√	√
61	P2_VIBTR02	-10	10	㎛	Shaft-vibration-related X-axis displacement near the 1st mass wheel	√ (VXT02)	√	√
62	P2_VIBTR03	-10	10	㎛	Shaft-vibration-related Y-axis displacement near the 2nd mass wheel	√ (VYT03)	√	√
63	P2_VIBTR04	-10	10	㎛	Shaft-vibration-related X-axis displacement near the 2nd mass wheel	√ (VXT03)	√	√
64	P2_VT01	11	12	rad/s	Phase lag signal of the key phasor probe	√	√	√
65	P2_VTR01	-10	10	㎛	Preset vibration limit for the sensor P2_VIBTR01		√	√
66	P2_VTR02	-10	10	㎛	Preset vibration limit for the sensor P2_VIBTR02		√	√
67	P2_VTR03	-10	10	㎛	Preset vibration limit for the sensor P2_VIBTR03		√	√
68	P2_VTR04	-10	10	㎛	Preset vibration limit for the sensor P2_VIBTR03		√	√
69	P3_FIT01	0	27,648	-	Flow rate of water flowing into the upper water tank		√	√
70	P3_LCP01D	0	27,648	-	Speed command for the pump LCP01	√	√	√
71	P3_LCV01D	0	27,648	-	Position command for the valve LCV01	√	√	√
72	P3_LH01	0	70	%	High water level set-point	√	√	√
73	P3_LIT01	0	90	%	Water level of the upper water tank	√ (LT01)	√	√
74	P3_LL01	0	70	%	Low water level set-point	√	√	√
75	P3_PIT01	0	27,648	-	Pressure of water flowing into the upper water tank		√	√
76	P4_HT_FD	-0.02	0.02	mHz	Frequency deviation of HTM	√	√	√
77	P4_HT_LD	0	100	MW	Electrical load demand of HTM	√	√	
78	P4_HT_PO	0	100	MW	Output power of HTM	√	√	√

No	Name	Range		Unit	Description	HAI		
		Min	Max			20.07	21.03	22.04 23.05
79	P4_HT_PS	0	100	MW	Scheduled power demand of HTM	√	√	√
80	P4_LD	0	500	MW	Total electrical load demand	√	√	√
81	P4_ST_FD	-0.02	0.02	Hz	Frequency deviation of STM	√	√	√
82	P4_ST_GOV	0	27,648	-	Gate opening rate of STM		√	√
83	P4_ST_LD	0	500	MW	Electrical load demand of STM	√	√	√
84	P4_ST_PO	0	500	MW	Output power of STM	√	√	√
85	P4_ST_PS	0	500	MW	Scheduled power demand of STM	√	√	√
86	P4_ST_PT01	0	27,648	-	Digital value of steam pressure of STM	√	√	√
87	P4_ST_TT01	0	27,648	-	Digital value of steam temperature of STM	√	√	√
TOTAL						59	78	86

HAIEND DATA POINTS

The HAIEnd 23.05 has 225 data points, all points of Boiler' Emerson Ovation DCS. 35 data points of HAIEnd correspond to the same points as in the HAI data points. Duplicate points are shown in the table below.

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
1	1001.02-OUT	-	-	-	Output of 1001.02	√	
2	1001.05-OUT	-	-	-	Output of 1001.05	√	
3	1001.07-OUT1	-	-	-	Output of 1001.07	√	
4	1001.07-OUT2	-	-	-	Output of 1001.07	√	
5	1001.08-OUT	-	-	-	Output of 1001.08	√	
6	1001.09-OUT	-	-	-	Output of 1001-.9	√	
7	1001.13-OUT	-	-	-	Output of 1001.13	√	
8	1001.14-OUT	-	-	-	Output of 1001.14	√	
9	1001.15-OUT	-	-	-	Output of 1001.15	√	√ (P1_B2016)
10	1001.16-OUT	-	-	-	Output of 1001.16	√	
11	1001.17-OUT	-	-	-	Output of 1001.17	√	
12	1001.20-OUT	-	-	-	Output of 1001.20	√	

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
13	1002.02-OUT	-	-	-	Output of 1002.02	√	
14	1002.06-OUT	-	-	-	Output of 1002.06	√	
15	1002.07-OUT	-	-	-	Output of 1002.07	√	
16	1002.08-OUT	-	-	-	Output of 1002.08	√	
17	1002.09-OUT	-	-	-	Output of 1002.09	√	
18	1002.11-OUT1	-	-	-	Output of 1002.11	√	
19	1002.11-OUT2	-	-	-	Output of 1002.11	√	
20	1002.12-OUT	-	-	-	Output of 1002.12	√	
21	1002.14-OUT	-	-	-	Output of 1002.14	√	
22	1002.15-OUT	-	-	-	Output of 1002.15	√	
23	1002.16-OUT1	-	-	-	Output of 1002.16	√	
24	1002.16-OUT2	-	-	-	Output of 1002.16	√	
25	1002.19-OUT	-	-	-	Output of 1002.19	√	
26	1002.20-OUT	-	-	-	Output of 1002.20	√	
27	1002.21-OUT	-	-	-	Output of 1002.21	√	
28	1002.29-OUT	-	-	-	Output of 1002.29	√	
29	1002.30-OUT	-	-	-	Output of 1002.30	√	
30	1002.31-OUT	-	-	-	Output of 1002.31	√	
31	1002.34-OUT	-	-	-	Output of 1002.34	√	
32	1003.05-OUT	-	-	-	Output of 1003.05	√	
33	1003.07-OUT	-	-	-	Output of 1003.07	√	
34	1003.10-OUT	-	-	-	Output of 1003.10	√	
35	1003.11-OUT	-	-	-	Output of 1003.11	√	
36	1003.12-OUT1	-	-	-	Output of 1003.12	√	
37	1003.12-OUT2	-	-	-	Output of 1003.12	√	
38	1003.13-OUT	-	-	-	Output of 1003.13	√	
39	1003.17-OUT	-	-	-	Output of 1003.17	√	
40	1003.18-OUT	-	-	-	Output of 1003.18	√	

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
41	1003.23-OUT	-	-	-	Output of 1003.23	√	
42	1003.24-OUT	-	-	-	Output of 1003.24	√	
43	1003.25-OUT	-	-	-	Output of 1003.25	√	
44	1003.26-OUT	-	-	-	Output of 1003.26	√	
45	1003.27-OUT	-	-	-	Output of 1003.27	√	
46	1003.29-OUT	-	-	-	Output of 1003.29	√	
47	1003.30-OUT	-	-	-	Output of 1003.30		
48	1004.11-OUT1	-	-	-	Output of 1004.11	√	
49	1004.11-OUT2	-	-	-	Output of 1004.11	√	
50	1004.12-OUT1	-	-	-	Output of 1004.12	√	
51	1004.12-OUT2	-	-	-	Output of 1004.12	√	
52	1004.13-OUT	-	-	-	Output of 1004.13	√	
53	1004.15-OUT1	-	-	-	Output of 1004.15	√	
54	1004.15-OUT2	-	-	-	Output of 1004.15	√	
55	1004.18-OUT1	-	-	-	Output of 1004.18	√	
56	1004.18-OUT2	-	-	-	Output of 1004.18	√	
57	1004.21-OUT	-	-	-	Output of 1004.21	√	
58	1004.24-OUT	-	-	-	Output of 1004.24	√	
59	1004.29-OUT	-	-	-	Output of 1004.29	√	
60	1004.36-OUT	-	-	-	Output of 1004.36	√	
61	1004.37-OUT	-	-	-	Output of 1004.37	√	
62	1004.38-OUT	-	-	-	Output of 1004.38	√	
63	1004.39-OUT	-	-	-	Output of 1004.39	√	
64	1004.41-OUT	-	-	-	Output of 1004.41	√	
65	1004.44-OUT	-	-	-	Output of 1004.44	√	
66	1004.52-OUT	-	-	-	Output of 1004.52	√	
67	1004.53-OUT	-	-	-	Output of 1004.53	√	
68	1004.62-OUT	-	-	-	Output of 1004.62	√	

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
69	1004.76-OUT	-	-	-	Output of 1004.76	√	
70	1004.78-OUT	-	-	-	Output of 1004.78	√	
71	1004.79-OUT	-	-	-	Output of 1004.79	√	
72	1004.80-OUT	-	-	-	Output of 1004.80	√	
73	1010.02-OUT	-	-	-	Output of 1010.02	√	
74	1010.03-OUT	-	-	-	Output of 1010.03	√	
75	1010.04-OUT	-	-	-	Output of 1010.04	√	
76	1010.05-OUT1	-	-	-	Output of 1010.05	√	
77	1010.05-OUT2	-	-	-	Output of 1010.05	√	
78	1010.05-OUT3	-	-	-	Output of 1010.05	√	
79	1010.05-OUT4	-	-	-	Output of 1010.05	√	
80	1010.07-OUT	-	-	-	Output of 1010.07	√	
81	1010.08-OUT	-	-	-	Output of 1010.08	√	
82	1010.09-OUT	-	-	-	Output of 1010.09	√	
83	1010.10-OUT	-	-	-	Output of 1010.10	√	
84	1010.11-OUT	-	-	-	Output of 1010.11	√	
85	1010.12-OUT1	-	-	-	Output of 1010.12	√	
86	1010.12-OUT2	-	-	-	Output of 1010.12	√	
87	1010.16-OUT	-	-	-	Output of 1010.16	√	
88	1010.17-OUT	-	-	-	Output of 1010.17	√	
89	1010.19-OUT	-	-	-	Output of 1010.19	√	
90	1010.23-OUT	-	-	-	Output of 1010.23	√	
91	1010.30-OUT	-	-	-	Output of 1010.30	√	
92	1010.31-OUT	-	-	-	Output of 1010.31	√	
93	1010.33-OUT	-	-	-	Output of 1010.33	√	
94	1010.35-OUT	-	-	-	Output of 1010.35	√	
95	1010.38-OUT	-	-	-	Output of 1010.38	√	
96	1010.39-OUT	-	-	-	Output of 1010.39	√	

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
97	1010.41-OUT	-	-	-	Output of 1010.41	√	
98	1010.42-OUT	-	-	-	Output of 1010.42	√	
99	1010.44-OUT	-	-	-	Output of 1010.44	√	
100	1010.46-OUT	-	-	-	Output of 1010.46	√	
101	1010.47-OUT	-	-	-	Output of 1010.47	√	
102	1010.50-OUT	-	-	-	Output of 1010.50	√	
103	1010.54-OUT	-	-	-	Output of 1010.54	√	
104	1010.56-OUT	-	-	-	Output of 1010.56	√	
105	1011.02-OUT	-	-	-	Output of 1011.02	√	
106	1011.03-OUT	-	-	-	Output of 1011.03	√	
107	1011.04-OUT	-	-	-	Output of 1011.04	√	
108	1011.05-OUT1	-	-	-	Output of 1011.05	√	
109	1011.05-OUT2	-	-	-	Output of 1011.05	√	
110	1011.05-OUT3	-	-	-	Output of 1011.05	√	
111	1011.05-OUT4	-	-	-	Output of 1011.05	√	
112	1011.07-OUT	-	-	-	Output of 1011.07	√	
113	1011.08-OUT	-	-	-	Output of 1011.08	√	
114	1011.09-OUT	-	-	-	Output of 1011.09	√	
115	1011.10-OUT	-	-	-	Output of 1011.10	√	
116	1011.11-OUT	-	-	-	Output of 1011.11	√	
117	1011.12-OUT1	-	-	-	Output of 1011.12	√	
118	1011.12-OUT2	-	-	-	Output of 1011.12	√	
119	1011.16-OUT	-	-	-	Output of 1011.16	√	
120	1011.17-OUT	-	-	-	Output of 1011.17	√	
121	1011.19-OUT	-	-	-	Output of 1011.19	√	
122	1011.23-OUT	-	-	-	Output of 1011.23	√	
123	1011.30-OUT	-	-	-	Output of 1011.30	√	
124	1011.31-OUT	-	-	-	Output of 1011.31	√	

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
125	1011.33-OUT	-	-	-	Output of 1011.33	√	
126	1011.35-OUT	-	-	-	Output of 1011.35	√	
127	1011.38-OUT	-	-	-	Output of 1011.38	√	
128	1011.39-OUT	-	-	-	Output of 1011.39	√	
129	1011.41-OUT	-	-	-	Output of 1011.41	√	
130	1011.42-OUT	-	-	-	Output of 1011.42	√	
131	1011.44-OUT	-	-	-	Output of 1011.44	√	
132	1011.46-OUT	-	-	-	Output of 1011.46	√	
133	1011.47-OUT	-	-	-	Output of 1011.47	√	
134	1011.50-OUT	-	-	-	Output of 1011.50	√	
135	1011.54-OUT	-	-	-	Output of 1011.54	√	
136	1011.56-OUT	-	-	-	Output of 1011.56	√	
137	1020.02-OUT	-	-	-	Output of 1020.02	√	
138	1020.03-OUT	-	-	-	Output of 1020.03	√	
139	1020.04-OUT	-	-	-	Output of 1020.04	√	
140	1020.05-OUT1	-	-	-	Output of 1020.05	√	
141	1020.05-OUT2	-	-	-	Output of 1020.05	√	
142	1020.09-OUT	-	-	-	Output of 1020.09	√	
143	1020.10-OUT	-	-	-	Output of 1020.10	√	
144	1020.11-OUT	-	-	-	Output of 1020.11	√	
145	1020.13-OUT	-	-	-	Output of 1020.13	√	
146	1020.14-OUT	-	-	-	Output of 1020.14	√	
147	1020.15-OUT	-	-	-	Output of 1020.15	√	
148	1020.18-OUT	-	-	-	Output of 1020.18	√	
149	1020.20-OUT	-	-	-	Output of 1020.20	√	
150	1020.21-OUT	-	-	-	Output of 1020.21	√	
151	DM-AIT-DO	-	-	-	For maintenance	√	
152	DM-AIT-PH	-	-	-	For maintenance	√	

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
153	DM-CIP-1ST	-	-	-	For maintenance	√	
154	DM-CIP-2ND	-	-	-	For maintenance	√	
155	DM-CIP-START	-	-	-	For maintenance	√	
156	DM-CIP-STEP1	-	-	-	For maintenance	√	
157	DM-CIP-STEP11	-	-	-	For maintenance	√	
158	DM-CIPH-1ST	-	-	-	For maintenance	√	
159	DM-CIPH-2ND	-	-	-	For maintenance	√	
160	DM-CIPH-START	-	-	-	For maintenance	√	
161	DM-CIPH-STEP1	-	-	-	For maintenance	√	
162	DM-CIPH-STEP11	-	-	-	For maintenance	√	
163	DM-COOL-ON	-	-	-	For maintenance	√	
164	DM-COOL-R	-	-	-	For maintenance	√	
165	DM-FCV01-D	0	100	%	Position command for the FCV01 valve	√	√ (P1_FCV01D)
166	DM-FCV01-Z	0	100	%	Current position of the FCV01 valve	√	√ (P1_FCV01Z)
167	DM-FCV02-D	0	100	%	Position command for the FCV02 valve	√	√ (P1_FCV02D)
168	DM-FCV02-Z	0	100	%	Current position of the FCV02 valve	√	√ (P1_FCV02Z)
169	DM-FCV03-D	0	100	%	Position command for the FCV03 valve	√	√ (P1_FCV03D)
170	DM-FCV03-Z	0	100	%	Current position of the FCV03 valve	√	√ (P1_FCV03Z)
171	DM-FT01	0	2,500	mmH2O	Measured flowrate of the return water tank	√	√ (P1_FT01)
172	DM-FT01Z	0	3,190	l/h	Water inflow rate converted from P1_FT01	√	√ (P1_FT01Z)
173	DM-FT02	0	2,500	mmH2O	Measured flowrate of heating water tank	√	√ (P1_FT02)
174	DM-FT02Z	0	3,190	l/h	Water outflow rate conversion from P1_FT02	√	√ (P1_FT02Z)
175	DM-FT03	0	2,500	mmH2O	Measured flowrate of the return water tank	√	√ (P1_FT03)
176	DM-FT03Z	0	3,190	l/h	Water outflow rate converted from P1_FT03	√	√ (P1_FT03Z)
177	DM-HT01-D	0	1	Boolean	Start command of heater	√	
178	DM-LCV01-D	0	100	%	Position command for the LCV01 valve	√	√ (P1_LCV01D)
179	DM-LCV01-MIS	0	1	Boolean	Check point when the difference between DM-LCV01-D and DM-LCV01-Z exceeds 10	√	
180	DM-LCV01-Z	0	100	%	Current position of the LCV01 valve	√	√ (P1_LCV01Z)

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
181	DM-LIT01	0	720	mm	Water level of the return water tank	√	√ (P1_LIT01)
182	DM-LSH-03	0	1	Boolean	Level high alarm of Return water tank (TK03)	√	
183	DM-LSH-04	0	1	Boolean	Level high alarm of Buffer water tank (TK04)	√	
184	DM-LSH01	0	1	Boolean	Level high alarm of Main water tank (TK01)	√	
185	DM-LSH02	0	1	Boolean	Level high alarm of Heating water tank (TK02)	√	
186	DM-LSL-04	0	1	Boolean	Level low alarm of Buffer water tank (TK04)	√	
187	DM-LSL01	0	1	Boolean	Level low alarm of Main water tank (TK01)	√	
188	DM-LSL02	0	1	Boolean	Level low alarm of Heating water tank (TK02)	√	
189	DM-PCV01-D	0	100	%	Position command for the PCV01 valve	√	√ (P1_PCV01D)
190	DM-PCV01-DEV	0	1	Boolean	Main pump discharge pressure control	√	
191	DM-PCV01-Z	0	100	%	Current position of the PCV01 valve	√	√ (P1_PCV01Z)
192	DM-PCV02-D	0	100	%	Position command for the PCV2 valve	√	√ (P1_PCV02D)
193	DM-PCV02-Z	0	100	%	Current position of the PCV02 valve	√	√ (P1_PCV02Z)
194	DM-PIT01-HH	0	10	bar	Highest outlet pressure of the heat-exchanger	√	√ (P1_PIT01_HH)
195	DM-PIT01	0	10	bar	Heat-exchanger outlet pressure	√	√ (P1_PIT01)
196	DM-PIT02	0	10	bar	Water supply pressure of the heating water pump	√	√ (P1_PIT02)
197	DM-PP01-R	0	1	Boolean	Running state of the boiler process	√	
198	DM-PP01A-D	0	1	Boolean	Start command of the main water pump PP01A	√	√ (P1_PP01AD)
199	DM-PP01A-R	0	1	Boolean	Running state of the main water pump PP01A	√	√ (P1_PP01AR)
200	DM-PP01B-D	0	1	Boolean	Start command of the main water pump PP01B	√	√ (P1_PP01BD)
201	DM-PP01B-R	0	1	Boolean	Running state of the main water pump PP01B	√	√ (P1_PP01BR)
202	DM-PP02-D	0	1	Boolean	Start command of the heating water pump PP02	√	√ (P1_PP02D)
203	DM-PP02-R	0	1	Boolean	Running state of the heating water pump PP02	√	√ (P1_PP02R)
204	DM-PP04-AO	0	100	Hz	Speed of the cooling water pump PP04	√	
205	DM-PP04-D	0	1	Boolean	Start command of the cooling water pump PP04	√	
206	DM-PP04-SV	-	-	-	For maintenance	√	
207	DM-PWIT-03	0	10	bar	Pressure of the heating water tank (TK02)	√	
208	DM-SOL01-D	0	1	Boolean	Open command of the main water tank (TK01) supply valve	√	√ (P1_SOL01D)

No	Name	Range		Unit	Description	HAIEnd	HAI
		Min	Max			23.05	23.05
209	DM-SOL02-D	0	1	Boolean	Open command of the return water tank (TK03) supply valve	√	
210	DM-SOL03-D	0	1	Boolean	Open command of the main water tank (TK01) drain valve	√	√ (P1_SOL03D)
211	DM-SOL04-D	0	1	Boolean	Open command of the return water tank (TK03) drain valve	√	
212	DM-SS01-RM	0	1	Boolean	Operation mode of the boiler DCS from control panel	√	
213	DM-ST-SP	0	1	Boolean	Start/stop command of the boiler DCS	√	√ (P1_STSP)
214	DM-SW01-ST	0	1	Boolean	Start command of the boiler DCS from control panel	√	
215	DM-SW02-SP	0	1	Boolean	Stop command of the boiler DCS from control panel	√	
216	DM-SW03-EM	0	1	Boolean	Emergency command of the boiler DCS local control panel	√	
217	DM-TIT01	-50	150	℃	Heat-exchanger outlet temperature	√	√ (P1_TIT01)
218	DM-TIT02	-50	150	℃	Temperature of the heating water tank	√	√ (P1_TIT03)
219	DM-TWIT-03	-50	150	℃	Temperature of the main water tank (TK01)	√	
220	DM-TWIT-04	-50	150	℃	Temperature of the return water tank (TK03)	√	
221	DM-TWIT-05	-50	150	℃	Temperature of the buffer water tank (TK04)	√	
222	DQ03-LCV01-D	0	1	Boolean	Alarm when measured value of LCV01 exceeds 50%	√	
223	DQ04-LCV01-DEV	0	1	Boolean	Point assignment from DQ04-LCV01-MIS	√	
224	GATEOPEN	0	10	-	Gate opening rate of STM	√	√ (P4_ST_GOV)
225	PP04-SP-OUT	0	40	℃	Running temperature of cooling	√	
TOTAL						225	35

ATTACK SCENARIOS

All attack scenarios in the viewpoint of a feedback control scheme were configured based on four types of variables, namely the setpoints (SPs), process variables (PVs), control variables (CVs), and control parameters (CPs). An attacker can control all variables by indirectly manipulating any algorithm blocks in the embedded controllers such as the setpoint algorithm, PID controller, signal conditioner and others. Thus, an attacker can ultimately achieve a stealthy attack on the control device.

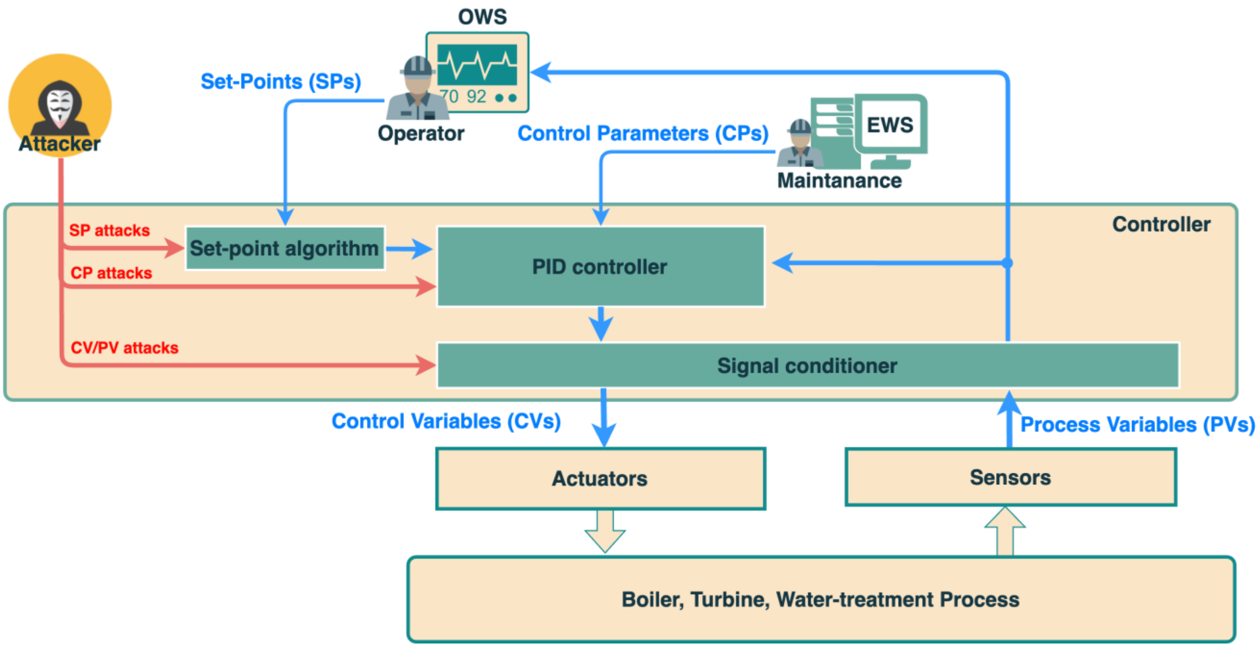

FIGURE 14. ATTACK MODEL BASED ON A PROCESS CONTROL LOOP(PCL).

The variables and parameters of the process control loop(PCL) are represented as "points" on the control device. A point is a variable allocated in memory for data access. I/O points are used for external peripheral device connection, whereas internal points are used for internal variables. The SP, PV, and CV variables of PCL correspond to I/O points. These points exchange information through hard-wired connections with external devices such as OWS and remote I/O. The internal points are only accessible to EWS that corresponds to CPs of PCL.

Normal Behaviors

During normal operation, it is assumed that the operator operates the control facility in a routine manner via the HMI, and that the simulator variables associated with power generation in the HIL simulator are changed. The operator monitors the PV values given by the current sensor displayed on the HMI, and adjusts the SPs of the various control devices to operate the system.

HMI operation task scheduler was used to periodically set the SPs and HIL simulator variables to random or predefined values within the normal range to simulate a benign scenario. The normal ranges of SP values in which the entire process was stable were determined by experimentally changing the value of each SP.

The four controllers (P1-PC, P1-LC, P1-FC, and P1-TC) and two simulation models (steam turbine power generator and pump-storage hydropower generator) were automatically operated several times a day. These were initiated with a random delay, and a random value or predefined value within the normal operational range was reached. All SP values were recorded to learn the system features.

Attack Behaviors

Attack scenarios are classified into the two categories depending on the attack points.

- I/O point: This type of attack indirectly manipulates SP, PV, and CO points, which are PCL parameters, through I/O point manipulation. This type of attack scenarios has been used to generate test data for all versions of HAI datasets.
- Internal point: This type of attack manipulates the parameter value of algorithm function that determines the internal points. Depending on function, this attack is activated when a specific condition is satisfied. This type of attack scenarios was only used in HAI/HAIEnd 23.05.

ATTACK SCENARIOS TARGETING I/O POINTS

Since 2019, attack scenarios targeting I/O points have been continuously developed, and the attack scenarios have been implemented by considering attack target, attack time, and method for each feedback control loop.

Scenario	Target			Description	HAI			
	Controller	Variable	Point		20.07	21.03	22.04	23.05
AP01	P1-PC	SP1	P1_B2016	Decrease or increase SP value of P1-PC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.	✓	✓	✓	✓
AP02	P1-PC	SP1	P1_B2016	Decrease or increase SP value of P1-PC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.	✓	✓	✓	✓
		PV1	P1_PIT01	Attempt to maintain previous sensor value.				
AP03	P1-PC	SP1	P1_B2016	Decrease or increase SP value of P1-PC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.			✓	✓
		PV1	P1_PIT01	Attempt to maintain previous sensor value.				
		PV2	P1_FIT01	Attempt to maintain previous sensor value				
AP04	P1-PC	CV1	P1_PCV01D	Decrease or increase CV value of P1-PC. Restore to normal.	✓	✓	✓	✓
AP05	P1-PC	CV1	P1_PCV01D	Decrease or increase CV value of P1-PC. Restore to normal.	✓	✓	✓	✓
		PV1	P1_PIT01	Attempt to maintain previous sensor value.				
AP06	P1-PC	SP1-ST	P1_B2016	Short-term (ST) attack that decrease or increase SP value of P1-PC for a few seconds and restores to normal. Repeat several times while hiding SP changes in HMI.		✓		
AP07	P1-PC	CV1-ST	P1_PCV01D	Short-term (ST) attack that decrease or increase CV value of P1-PC for a few seconds and restores to normal. Repeat several times while hiding SP changes in HMI.			✓	✓
AP08	P1-FC	SP1	P1_B3005	Decrease or increase SP value of P1-FC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI	✓	✓	✓	✓
AP09	P1-FC	SP1	P1_B3005	Decrease or increase SP value of P1-FC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI	✓	✓	✓	✓
		PV1	P1_FT03	Attempt to maintain previous sensor value.				

Scenario	Target			Description	HAI			
	Controller	Variable	Point		20.07	21.03	22.04	23.05
AP10	P1-FC	SP1	P1_B3005	Decrease or increase SP value of P1-FC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI			√	√
		PV1	P1_FT03	Attempt to maintain previous sensor value.				
		PV2	P1_LIT01	Attempt to maintain previous sensor value.				
AP11	P1-FC	CV1	P1_FCV03D	Decrease or increase CV value of P1-FC. Restore in form of trapezoidal profile.		√	√	√
AP12	P1-FC	CV1	P1_FCV03D	Decrease or increase CV value of P1-FC. Restore to normal.		√	√	√
		PV1	P1_FT03	Attempt to maintain previous sensor value.				
AP13	P1-FC	CV1-ST	P1_FCV03D	Short-term (ST) attack that decrease or increase CV value of P1-FC for a few seconds and restores to normal. Repeat several times while hiding SP changes in HMI.		√	√	√
AP14	P1-LC	SP1	P1_B3004	Decrease or increase SP value of P1-LC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.	√	√	√	√
AP15	P1-LC	SP1	P1_B3004	Decrease or increase SP value of P1-LC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.	√	√	√	√
		PV1	P1_LIT01	Attempt to repeat previous sensor value.				
AP16	P1-LC	CV1	P1_LCV01D	Decrease or increase CV value of P1-LC. Restore to normal.	√	√	√	√
AP17	P1-LC	CV1	P1_LCV01D	Decrease or increase CV value of P1-LC. Restore to normal.	√	√	√	√
		PV1	P1_LIT01	Attempt to repeat previous sensor value.				
AP18	P1-LC	CV1-ST	P1_LCV01D	Short-term (ST) attack that decrease or increase CV value of P1-LC for a few seconds and restores to normal. Repeat several times while hiding SP changes in HMI.		√	√	√
AP19	P1-TC	CV1	P1_FCV01D	Decrease or increase CV value of P1-TC. Restore to normal.			√	√
AP20	P1-TC	CV1	P1_FCV01D	Decrease or increase CV value of P1-TC. Restore to normal.			√	√
		PV1	P1_TIT01	Attempt to repeat previous sensor value.				
AP21	P1-TC	CV1-ST	P1_FCV01D	Short-term (ST) attack that decrease or increase CV value of P1-TC for a few seconds and restores to normal. Repeat several times while hiding SP changes in HMI.			√	√
AP22	P1-TC	SP1-LT	P1_B4002	Long-term (LT) attack that decrease or increase SP value of P1-TC continuously for more than 10 minutes and restores to normal.			√	√
AP23	P1-CC	CV1	P1_PP04	Decrease or increase CV value of P1-CC. Restore to normal.			√	√
AP24	P1-CC	CV1-ST	P1_PP04	Short-term (ST) attack that decrease or increase CV value of P1-CC for a few seconds			√	√

Scenario	Target			Description	HAI			
	Controller	Variable	Point		20.07	21.03	22.04	23.05
				and restores to normal. Repeat several times while hiding SP changes in HMI.				
AP25	P1-CC	SP1-LT	P1_PP04_SP	Long-term (LT) attack that decrease or increase SP value of P1-CC continuously for more than 10 minutes and restores to normal.			√	√
AP26	P2-SC	SP1	P2_AutoSD (P2_SD01)	Decrease or increase SP value of P2-SC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.	√	√	√	√
AP27	P2-SC	SP1	P2_AutoSD (P2_SD01)	Decrease or increase SP value of P2-SC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.	√	√	√	√
		PV1	P2_SIT01	Attempt to maintain previous sensor value.				
AP28	P2-SC	SP2	P2_ManualSD	Decrease or increase SP value of P2-SC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.			√	
AP29	P2-SC	CV1	P2_SCO	Decrease or increase CV value of P2-SC. Restore to normal.		√	√	
AP30	P2-SC	CV1	P2_SCO	Decrease or increase CV value of P2-SC. Restore to normal.		√	√	√
		PV1	P2_SIT01	Attempt to maintain previous sensor value.				
AP31	P2-SC	SP1-ST	P2_AutoSD	Short-term (ST) attack that decrease or increase CV value of P2-SC for a few seconds and restores to normal. Repeat several times while hiding SP changes in HMI.		√	√	
AP32	P2-TC	SP1	P2_VTR01	Decrease or increase SP value of P2-TC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.		√		√
AP33	P2-TC	SP2	P2_VTR02	Decrease or increase SP value of P2-SC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.		√	√	
AP34	P2-TC	SP3	P2_RTR	Decrease or increase SP value of P2-SC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.		√	√	
AP35	P3-LC	CV1	P3_LCP01D	Attempt to repeat previous sensor value.	√	√	√	√
AP36	P3-LC	CV1	P3_LCP01D	Decrease or increase CV value of P3-LC. Restore to normal.			√	
		PV1	P3_LIT01	Attempt to maintain previous sensor value.				
AP37	P3-LC	CV2	P3_LCV01D	Decrease or increase CV value of P3-LC. Restore to normal.	√	√	√	
AP38	P3-LC	CV2	P3_LCV01D	Decrease or increase CV value of P3-LC. Restore to normal.			√	
		PV1	P3_LIT01	Attempt to maintain previous sensor value.				
AP39	P3-LC	CV2-LT	P3_LCV01D	Long-term (LT) attack that decrease or increase CV value of P3-LC continuously for more than 10 minutes and restores to normal.			√	
AP40	P1-PC	SP1-LT	P1_B2016	Long-term (LT) attack that decrease or increase SP value of P1-PC continuously for more than 10 minutes and restores to normal.				√

Scenario	Target			Description	HAI			
	Controller	Variable	Point		20.07	21.03	22.04	23.05
AP41	P1-FC	SP1-LT	P1_B3005	Long-term (LT) attack that decrease or increase SP value of P1-FC continuously for more than 10 minutes and restores to normal.				√
AP42	P1-LC	CV1	P1_LCV01D	Decrease or increase CV value of P1-LC. Restore to normal.				√
		PV1	P1_LIT01	Attempt to repeat previous sensor value.				
		PV2	P1_FT03	Attempt to maintain previous sensor value.				
AP43	P1-LC	CV1-LT	P1_LCV01D	Long-term (LT) attack that decrease or increase CV value of P1-LC continuously for more than 10 minutes and restores to normal.				√
AP44	P1-LC	CV1-LT	P1_LCV01D	Decrease or increase CV value of P1-LC. Restore to normal.				√
		PV1-LT	P1_LIT01	Attempt to repeat previous sensor value.				
AP45	P1-TC	SP1	P1_B4002	Decrease or increase SP value of P1-TC. Restore as a form of a trapezoidal profile while hiding SP changes in HMI.				√
AP46	P1-CC	CV1	P1_PP04	Decrease or increase CV value of P1-CC. Restore to normal.				√
		PV1	P1_TIT03	Attempt to repeat previous sensor value.				
AP47	P2-TC	SP2-LT	P2_VTR02	Long-term (LT) attack that decrease or increase SP value of P2-TC continuously for more than 10 minutes and restores to normal.				√
TOTAL					14	25	37	37

ATTACK SCENARIOS TARGETING INTERNAL POINTS

Attack scenarios targeting internal points modulate the algorithm function used inside the control logic. The boiler DCS of the HAI testbed employs various algorithm functions, and we developed attack scenarios that target the artificial I/O, arithmetic, and monitor functions.

- Artificial I/O function: This function is used to initialize an algorithm function with internal parameters. An attacker can tamper the output of the algorithm function when the process is initialized by modulating the internal parameters.
- Arithmetic function: This function is utilized to generate calibration curves for sensor inputs or control command outputs. An attacker can degrade the performance of a sensor or controller by changing calibration tuning parameters.
- Monitor function: This function monitors whether the input signal crosses the high/low threshold. An attacker can change the threshold to cause over-detection or no-detection of anomalies.

A trigger-type attack, in which an attack only occurs in a specific situation, can be implemented if the internal parameters of the algorithm function are changed. The artificial I/O function can modulate the output of the control signal associated with the algorithm function into an arbitrary value at initialization. The arithmetic function triggers an attack only when the input value reaches the area affected by the calibration adjustment parameter. The monitor function triggers an attack only when the input value reaches the modulated high/low threshold. The eight attack scenarios are only used to generate HAI/HAIEnd 23.05; and are not used for previous version.

Scenario	Target			Description	HAI	HAIEnd
	Controller	Function	Attack Appeared Point		23.05	
AE01	P1-PC	1001.09	1001.09-OUT	Targets the algorithm, one of the artificial I/O functions. Modulate internal point(1001.09.R1) from 40(normal initial value) to 12. When the process changed from manual mode to auto mode, the modulated initialization value is sent to PCV01.	√	
AE02	P1-PC	1001.21	DM-PCV01-D	Targets the algorithm, one of the arithmetic functions; It scales by converting the input value into a one-to-one linear function. By modulating internal point(1001.21.S5) from 100(normal initial value) to 90, the maximum of PCV01 control command is modulated to 90 and decreases linearly.	√	
AE03	P1-LC	1002.14	1002.14-OUT	Targets the algorithm, one of the artificial I/O functions. Modulate internal point(1002.14.R1) from 30(normal initial value) to 0. When the process changed from manual mode to auto mode, the modulated initialization value is sent to LCV01.	√	
AE04	P1-LC	1002.31	1002.31-OUT	Targets the algorithm, one of an arithmetic functions; By modulating internal point(1002.31.S5) from 97(normal initial value) to 87, the maximum of LCV01 control command is modulated to 87 and decreases linearly.	√	
AE05	P1-TC	1003.05	1003.05-OUT	Targets the algorithm, one of an arithmetic functions; By modulating internal point(1003.05.S5) from 100(normal initial value) to 90, the maximum of PV(1003.26) is modulated to 90 and decreases linearly.	√	
AE06	P1-TC	1003.08	DM-FT02Z	Targets the algorithm, one of an arithmetic functions; By modulating internal point(1003.08.T6) from 3190(normal initial value) to 3000, the maximum of DM-FT02Z(flow rate of boiler hot water) is modulated to 3000 and decreases linearly.	√	
AE07	P1-CC	1020.15	1020.15-OUT	Targets the algorithm, one of an arithmetic functions; By modulating internal point(1020.15.S4) from 0(normal initial value) to 15, the minimum of PP04 control command is modulated to 15 and increases linearly.	√	
AE08	P1-HC	1004.21	1004.21-OUT	Targets the algorithm, one of a monitor functions; By decreasing internal point(1004.21.R1) from 33(normal initial value) to 10 and increasing it to 40, the Low threshold of heater is modulated.	√	
TOTAL					8	

DATASETS

Since 2020, four versions of the dataset have been released, and herein, these datasets are described in detail starting with latest version. It is noteworthy that the version numbering follows a date-based scheme, where the version number indicates the released date.

HAI/HAIEnd 23.05

HAI 23.05 and HAIEnd 23.05 were collected at the same time. HAI 23.05 and HAIEnd 23.05 include four training datasets, two testing datasets, and one label dataset in the form of CSV file. The time-series data in each CSV file satisfies time continuity. The first column represents the observed time in the “yyyy-MM-dd hh:mm:ss” format, and the remaining columns provide the recorded SCADA data points. The label dataset was marked as 1 only when attack occurred to indicate the presence or absence of an attack.

NORMAL OPERATION

We used a hidden Markov model (HMM) to model the normal operation of SCADA. The HMM probabilistically determines the sequence and the delivery time of set point commands from a set of seven set points. Three HMMs are constructed to generate normal operations of three process controllers of the HAI testbed. The internal states and transition probability were constructed by considering the general process of each process control. The set-points are finally output probabilistically as possible observations. The probabilistic parameters of all the HMMs were given below. The change value of each observation was randomly determined within its normal range.

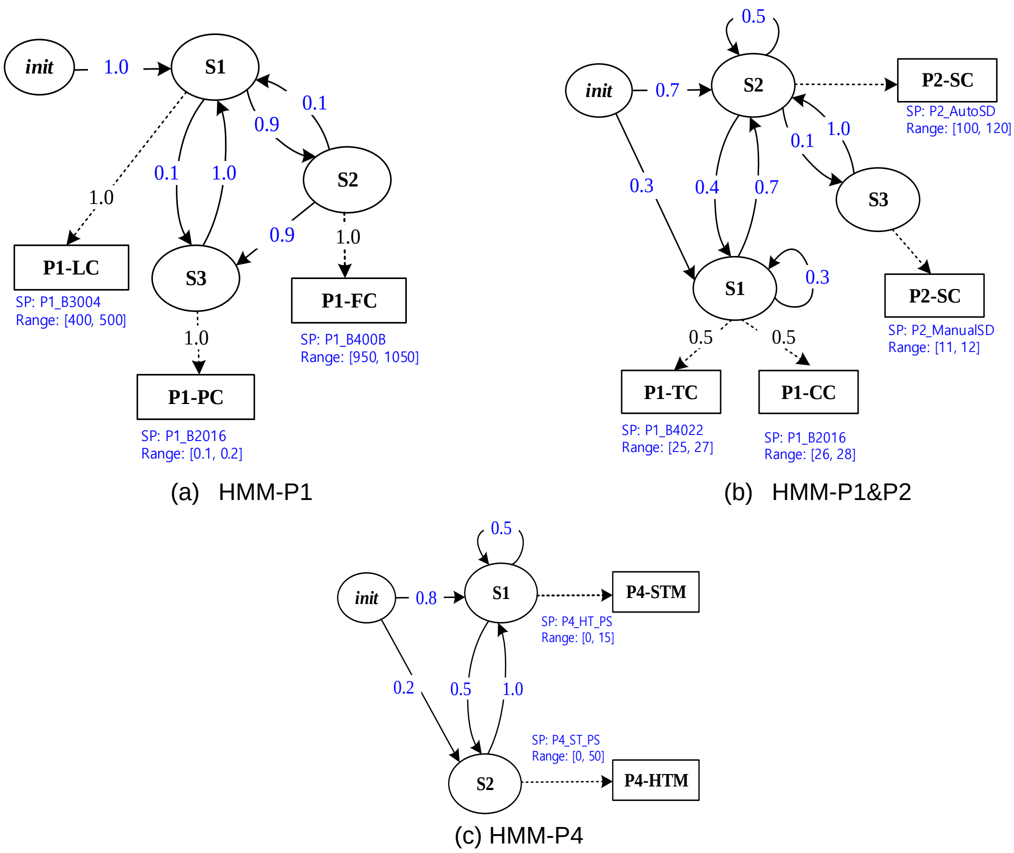

FIGURE 15. HMM-BASED GENERATIVE MODELS FOR NORMAL OPERATION.

ATTACK OPERATION

Because the collection target of HAIEnd is limited to the DCS of the boiler control system, the attack scenario also was implemented targeting the boiler control system.

The 52 attacks were conducted, including 42 attack primitives and 10 combinations of attacks designed to simultaneously perform two attack primitives. The attack scenarios are given below.

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
1	A101	AP01	P1-PC-SP1	P1_B2016	Aug. 12, 2022	16:25	237
2	A102	AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01		17:35	198
3	A103	AP04	P1-PC-CO1	P1_PCV01D		18:32	156
4	A104	AP05	P1-PC-CO1PV1	P1_PCV01D, P1_PIT01		19:21	164
5	A105	AP07	P1-PC-CO1-ST	P1_PCV01D		20:43	161
6	A106	AP03	P1-PC-SP1PV1PV2	P1_B2016, P1_PIT01		21:36	197
7	A107	AP40	P1-PC-SP1-LT	P1_B2016		22:47	604
8	A108	AP08	P1-FC-SP1	P1_B3005		23:35	96
9	A109	AP09	P1-FC-SP1PV1	P1_B3005, P1_FT03	Aug. 13, 2022	0:25	130
10	A110	AP11	P1-FC-CO1	P1_FCV03D		1:34	55
11	A111	AP12	P1-FC-CO1PV1	P1_FCV03D, P1_FT03		2:21	131
12	A112	AP13	P1-FC-CO1-ST	P1_FCV03D		3:26	78
13	A113	AP10	P1-FC-SP1PV1PV2	P1_B3005, P1_FT03, P1_LIT01		4:43	133
14	A114	AP41	P1-FC-SP1-LT	P1_B3005		5:40	627
15	A201	AP14	P1-LC-SP1	P1_B3004	Aug. 17, 2022	1:27	132
16	A202	AP15	P1-LC-SP1PV1	P1_B3004, P1_LIT01		3:37	131
17	A203	AP16	P1-LC-CO1	P1_LCV01D		4:21	68
18	A204	AP17	P1-LC-CO1PV1	P1_LCV01D, P1_LIT01		5:46	122
19	A205	AP18	P1-LC-CO1-ST	P1_LCV01D		6:21	85
20	A206	AP03	P1-PC-SP1PV1PV2	P1_B2016, P1_PIT01, P1_FIT		8:36	196
21	A207	AP43	P1-LC-CO1-LT	P1_LCV01D		9:42	614
22	A208	AP42	P1-LC-CO1PV1PV2	P1_LCV01D, P1_LIT01, P1_FIT		10:36	133
23	A209	AP23	P1-CC-CO1	P1_PP04		11:35	85
24	A210	AP23	P1-CC-CO1	P1_PP04		12:25	88
25	A211	AP46	P1-CC-CO1PV1	P1_PP04, P1_TIT03		13:47	204
26	A212	AP24	P1-CC-CO1-ST	P1_PP04		14:25	127
27	A213	AP25	P1-CC-SP1-LT	P1_PP04_SP		15:13	539
28	A214	AP19	P1-TC-CO1	P1_FCV01D		17:34	61
29	A215	AP20	P1-TC-CO1PV1	P1_FCV01D, P1_TIT01		18:16	147
30	A216	AP21	P1-TC-CO1-ST	P1_FCV01D		19:40	95

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
31	A217	AP22	P1-TC-SP1-LT	P1_B4002		20:12	505
32	A218	AP07	P1-PC-CO1-ST	P1_PCV01D		22:41	214
33	A219	AP13	P1-FC-CO1-ST	P1_FCV03D		23:38	131
34	A220	AE03	P1-LC-CO-AVALGEN	1002-14-AVALGEN.R1		13:48	131
35	A221	AE08	P1-HC-SP-LOWMON	1004-21-LOWMON.R1		14:58	82
36	A222	AE01	P1-PC-CO-AVALGEN	1001-09-AVALGEN.R1	Aug. 18, 2022	16:20	211
37	A223	AE07	P1-CC-CO-FUNCTION	1020-15-FUNCTION.S4		17:38	79
38	A224	AP14	P1-LC-SP1	P1_B3004		18:45	107
		AP26	P2-SC-SP1	P2_AutoSD			
39	A225	AP16	P1-LC-CO1	P1_LCV01D		19:21	60
		AP32	P2-TC-SP1	P2_VTR01			
40	A226	AP04	P1-PC-CO1	P1_PCV01D		20:32	118
		AP11	P1-FC-CO1	P1_FCV03D			
41	A227	AP09	P1-FC-SP1PV1	P1_B3005, P1_FT03		21:41	132
		AP14	P1-LC-SP1	P1_B3004			
42	A228	AP05	P1-PC-CO1PV1	P1_PCV01D, P1_PIT01		23:15	155
		AP30	P2-SC-CO1PV1	P2_SCO, P2_SIT01			
43	A229	AP45	P1-TC-SP1	P1_B4002	Aug. 19, 2022	1:23	115
		AP01	P1-PC-SP1	P1_B2016			
44	A230	AP19	P1-TC-CO1	P1_FCV01D		2:43	154
		AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01			
45	A231	AP08	P1-FC-SP1	P1_B3005		4:34	95
		AP35	P3-LC-CO1	P3_LCP01D			
46	A232	AP45	P1-TC-SP1	P1_B4002		5:14	153
		AP27	P2-SC-SP1PV1	P2_AutoSD, P2_SIT01			
47	A233	AP44	P1-LC-CO1PV1-LT	P1_LCV01D, P1_LIT01		6:46	2051
		AP47	P2-TC-SP2-LT	P2_VTR02			
48	A234	AP25	P1-CC-SP1-LT	P1_PP04_SP		8:24	529
49	A235	AE05	P1-TC-PV_FUNCTION1	1003-05-FUNCTION.T9		9:27	86
50	A236	AE06	P1-TC-PV_FUNCTION2	OCB0004018.T6		10:34	119
51	A237	AE05	P1-PC-CO_FUNCTION	OCB0002006.S5		14:18	189
52	A238	AE06	P1-LC-CO_FUNCTION	1002-31-FUNCTION.S5		14:51	122

HAI 22.04

HAI 22.04 includes six CSV files as training datasets and four CSV files as testing datasets. The time-series data in each CSV file satisfies time continuity and includes 89 columns. The first column represents the observed time in the “yyyy-MM-dd hh:mm:ss” format, while the next 87 columns provide the recorded SCADA data points. The last four columns provide data labels for the presence or absence of an attack. Out of these columns, the attack column is applicable to all processes and the other three columns are applicable to the corresponding control processes.

NORMAL OPERATION

We used a hidden Markov model (HMM) to model the normal operation of SCADA. The HMM probabilistically determines the sequence and the delivery time of set point commands from a set of seven set points. The probabilistic parameters of all the HMMs are the same as HAI 23.05.

ATTACK OPERATION

The 58 attacks were conducted, including 32 attack primitives and 26 combinations of attacks designed to simultaneously perform two attack primitives. The attack scenarios are given below.

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
1	A101	AP04	P1-PC-CO1	P1_PCV01D	Jul. 10, 2021	5:41	190
2	A102	AP18	P1-LC-CO1-ST	P1_LCV01D		7:19	54
3	A103	AP11	P1-FC-CO1PV1	P1_FCV03D, P1_FT03		11:25	126
4	A104	AP37	P3-LC-CO2	P3_LCV01D		15:39	54
5	A105	AP14	P1-LC-SP1	P1_B3004		16:42	296
6	A106	AP13	P1-CC-CO1	P1_PP04		19:21	91
7	A107	AP19	P1-TC-CO1	P1_FCV01D		22:35	67
8	A201	AP01	P1-PC-SP1	P1_B2016	Jul. 13, 2021	16:38	257
9	A202	AP13	P1-FC-CO1-ST	P1_FCV03D		17:21	65
10	A203	AP31	P2-SC-SP1-ST	P2_AutoSD		18:13	45
11	A204	AP04	P1-PC-CO1	P1_PCV01D		20:28	248
		AP29	P2-SC-CO1	P2_SCO			
12	A205	AP37	P3-LC-CO2	P3_LCV01D		21:10	55
13	A206	AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01		21:58	176
		AP27	P2-SC-SP1PV1	P2_AutoSD, P2_SIT01			
14	A207	AP16	P1-LC-CO1	P1_LCV01D		23:40	284
15	A208	AP30	P2-SC-CO1PV1	P2_SCO, P2_SIT01	Jul. 14, 2021	1:15	152
16	A209	AP03	P1-PC-SP1PV1PV2	P1_B2016, P1_PIT01, P1_FIT01		1:40	162
17	A210	AP26	P2-SC-SP1	P2_AutoSD		3:23	97
18	A211	AP05	P1-PC- CO1PV1	P1_PCV01D, P1_PIT01		7:21	151

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
19	A212	AP35	P3-LC-CO1	P3_LCP01D		8:11	55
20	A213	AP24	P1-CC-CO1-ST	P1_PP04		10:35	80
21	A214	AP39	P3-LC-CO2-LT	P3_LCV01D		11:23	613
22	A215	AP09	P1-FC-SP1PV1	P1_B3005, P1_FT03		12:17	168
23	A216	AP01	P1-PC-SP1	P1_B2016		13:52	158
		AP08	P1-FC-SP1	P1_B3005			
24	A217	AP10	P1-FC-CO1	P1_FCV03D		14:31	98
25	A301	AP16	P3-LC-CO2	P2_LCV01D		18:21	348
		AP10	P1-FC-CO1	P1_FCV03D			
26	A302	AP15	P1-LC-SP1PV1	P1_LCV01D		20:16	358
27	A303	AP17	P1-LC-CO1PV1	P1_B3004. P1_LIT01		23:22	143
		AP37	P3-LC-CO2	P3_LCV01D			
28	A304	AP38	P3-LC-CO2PV1	P1_LCV01D. P1_LIT01	Jul. 15, 2021	1:41	91
29	A305	AP18	P1-LC-CO1-ST	P3_LCV01D		2:09	94
30	A306	AP04	P1-PC-CO1	P1_LCV01D		3:37	353
		AP15	P1-LC-SP1PV1	P1_B3004. P1_LIT01			
31	A307	AP20	P1-TC-CO1PV1	P1_FCV01D. P1_TIT01		5:35	151
32	A308	AP05	P1-PC-CO1PV1	P1_PCV01D, P1_PIT01		6:53	173
		AP23	P1-CC-CO1	P1_PP04			
33	A309	AP08	P1-FC-SP1	P1_B3005		7:42	96
		AP19	P1-TC-CO1	P1_FCV01D			
34	A310	AP35	P3-LC-CO1	P3_LCP01D		9:52	2024
		AP37	P3-LC-CO2	P3_LCV01D			
35	A401	AP28	P2-SC-SP2	P2_ManualSD		12:42	38
36	A402	AP21	P1-TC-CO1-ST	P1_FCV01D		13:20	88
37	A403	AP34	P2-TC-SP3	P2_RTR		13:57	96
38	A404	AP26	P2-SC-SP1	P2_AutoSD		15:08	97
		AP37	P3-LC-CO2	P3_LCV01D			
39	A405	AP22	P1-TC-SP1-LT	P1_B4002		16:07	505
40	A406	AP09	P1-FC-SP1PV1	P1_B3005, P1_FT03		17:22	186
		AP19	P1-TC-CO1	P1_FCV01D			
41	A407	AP13	P1-FC-CO1-ST	P1_FCV03D		19:45	122

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
		AP17	P1-LC-CO1PV1	P1_LCV01D, P1_LIT01			
42	A408	AP05	P1-PC-CO1PV1	P1_PCV01D, P1_PIT01		20:29	673
		AP17	P1-LC-CO1PV1	P1_LCV01D, P1_LIT01			
43	A409	AP18	P1-LC-CO1-ST8	P1_LCV01D		22:41	63
		AP21	P1-TC-CO1-ST9	P1_FCV01D			
44	A410	AP11	P1-FC-CO1PV1	P1_FCV03D, P1_FT03		01:07	179
		AP27	P2-SC-SP1PV1	P2_AutoSD, P2_SIT01			
45	A411	AP23	P1-CC-CO1	P1_PP04		03:35	99
		AP34	P2-TC-SP3	P2_RTR			
	A412	AP20	P1-TC-CO1PV1	P1_FCV01D, P1_TIT01		04:02	156
		AP01	P1-PC-SP1	P1_B2016			
47	A413	AP16	P1-LC-CO1	P1_LCV01D		04:59	153
		AP27	P2-SC-SP1PV1	P2_AutoSD, P2_SIT01			
48	A414	AP33	P2-TC-SP2	P2_VTR02		07:20	77
		AP36	P3-LC-CO1PV1	P3_LCP01D, P3_LIT01			
49	A415	AP3	P2-TC-SP2	P2_VTR02		09:17	77
50	A416	AP12	P1-FC-CO1PV1PV2	P1_FCV03D, P1_FT03, P1_LIT01		10:39	134
51	A417	AP25	P1-CC-SP1-LT	P1_PP04_SP.	Jul. 16, 2021	11:22	544
52	A418	AP01	P1-PC-SP1	P1_B2016		13:23	342
		AP14	P1-LC-SP1	P1_B3004			
53	A419	AP01	P1-PC-SP1	P1_B2016		14:59	163
		AP35	P3-LC-CO1	P3_LCP01D			
54	A420	AP07	P1-PC-CO1-ST	P1_PCV01D		15:57	89
55	A421	AP30	P2-SC-CO1PV1	P2_SCO, P2_SIT01		17:34	152
		AP23	P1-CC-CO1	P1_PP04			
56	A422	AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01		20:08	165
		AP26	P2-SC-SP1	P2_AutoSD			
57	A423	AP08	P1-FC-SP1	P1_B3005		22:17	115
		AP29	P2-SC-CO1	P2_SCO			
58	A424	AP10	P1-FC-CO1	P1_FCV03D		23:05	86
		AP23	P1-CC-CO1	P1_PP04			

HAI 21.03

HAI 21.03 includes three CSV files as training datasets and five CSV files as testing datasets. The time-series data in each CSV file satisfies time continuity, and includes 84 columns. The first column represents the observed time as “yyyy-MM-dd hh:mm:ss,” while the next 78 columns provide the recorded SCADA data points. The last four columns provide data labels for whether an attack occurred or not, where the attack column was applicable to all process and the other three columns were for the corresponding control processes.

NORMAL OPERATION

An HMI operation task scheduler periodically sets the SPs and HIL simulator variables to predefined values within the normal range to simulate a benign scenario. The benign scenarios are given below.

No	Set points						Start Time
	P1_B2004 (Pressure SP)	P1_B3004 (Level SP)	P1_B3005 (Flowrate SP)	P1_B4002 (Temperature SP)	P4_ST_PS (Scheduled Power)	P4_HT_PS (Scheduled Power)	
1	0.1 (±0.002)	440 (±9)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	03:00 (±10)
2	0.03 (±0.001)	400 (±8)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	04:30 (±10)
3	0.1 (±0.002)	400 (±8)	1,100 (±22)	32 (±1)	0 (±0)	0 (±0)	06:00 (±10)
4	0.1 (±0.002)	400 (±8)	900 (±18)	32 (±0)	0 (±0)	0 (±0)	08:30 (±10)
5	0.1 (±0.002)	380 (±8)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	10:00 (±10)
6	0.06 (±0.001)	420 (±8)	1,000 (±20)	32 (±0)	0 (±0)	0 (±0)	12:00 (±0)
7	0.1 (±0.002)	400 (±40)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	14:30 (±10)
8	0.1 (±0.002)	400 (±8)	1,000 (±60)	33 (±1)	0 (±0)	0 (±0)	17:00 (±10)
9	0.1 (±0.002)	400 (±8)	1,100 (±22)	32 (±1)	0 (±0)	0 (±0)	19:30 (±10)
10	0.1 (±0.002)	400 (±8)	1,100 (±22)	32 (±1)	50 (±0)	10 (±0)	22:00 (±10)

ATTACK OPERATION

The 50 attacks were conducted, including 25 attack primitives and 25 combinations of attacks designed to simultaneously perform two attack primitives. The attack scenarios are given below.

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
1	**A101**	AP01	P1-PC-SP1	P1_B2016	Jul. 7, 2020	15:35	192
2	**A102**	AP06	P1-FC-SP1	P1_B3005		17:28	98
3	**A103**	AP13	P1-LC-CO1	P1_LCV01D		18:59	190
4	**A104**	AP18	P2-SC-CO1	P2_SCO		20:21	60
5	**A105**	AP16	P2-SC-SP1	P2_AutoSD		21:03	89
6	**A201**	AP22	P2-TC-SP2	P2_VTR02	Jul. 9, 2020	15:47	83
7	**A202**	AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01		17:38	422
8	**A203**	AP15	P1-LC-CO1-ST7	P1_LCV01D		18:59	17

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
9	A204	AP07	P1-FC-SP1PV1	P1_B3005, P1_FT03		20:10	259
10	A205	AP05	P1-PC-SP1-ST10	P1_B2016		21:15	123
11	A206	AP09	P1-FC-CO1PV1	P1_FCV03D, P1_FT03		23:02	256
12	A207	AP21	P2-TC-SP1	P2_VTR01	Jul. 10, 2020	01:08	68
13	A208	AP12	P1-LC-SP1PV1	P1_B3004, P1_LIT01		01:33	261
14	A209	AP11	P1-LC-SP1	P1_B3004		03:03	159
15	A210	AP04	P1-PC-CO1PV1	P1_PCV01D, P1_PIT01		05:29	421
16	A211	AP20	P2-SC-SP1-ST5	P2_AutoSD		07:51	45
17	A212	AP17	P2-SC-SP1PV1	P2_AutoSD, P2_SIT01		09:13	152
18	A213	AP14	P1-LC-CO1PV1	P1_LCV01D, P1_LIT01		10:49	254
19	A214	AP03	P1-PC-CO1	P1_PCV01D		12:51	152
20	A215	AP19	P2-SC-CO1PV1	P2_SCO, P2_SIT01		15:11	151
21	A216	AP10	P1-FC-CO1-ST10	P1_FCV03D		15:40	65
22	A217	AP23	P2-TC-SP3	P2_RTR		16:22	184
23	A218	AP08	P1-FC-CO1	P1_FCV03D		18:21	99
24	A219	AP24	P3-LC-CO1	P3_LCP01D		21:25	119
25	A220	AP25	P3-LC-CO2	P2_LCV01D		22:56	119
26	A301	AP15	P1-LC-CO1-ST	P1_LCV01D	Jul. 13, 2020	13:51	132
		AP06	P1-FC-SP1	P1_B3005			
27	A302	AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01		15:21	421
		AP06	P1-FC-SP1	P1_B3005			
28	A303	AP03	P1-PC-CO1	P1_PCV01D		18:11	189
		AP13	P1-LC-CO1	P1_LCV01D			
29	A304	AP16	P2-SC-SP1	P2_AutoSD		20:53	106
		AP21	P2-TC-SP1	P2_VTR01			
30	A305	AP18	P2-SC-CO1	P2_SCO		21:23	84
		AP22	P2-TC-SP2	P2_VTR02			
31	A306	AP01	P1-PC-SP1	P1_B2016		23:55	238
		AP16	P2-SC-SP1	P2_AutoSD			
32	A307	AP08	P1-FC-CO1	P1_FCV03D	Jul. 14, 2020	01:51	110
		AP21	P2-TC-SP1	P2_VTR01			
33	A308	AP14	P1-LC-CO1PV1	P1_LCV01D, P1_LIT01		03:53	255
		AP20	P2-SC-SP1-ST	P2_AutoSD			
34	A401	AP03	P1-PC-CO1	P1_PCV01D	Jul. 28,	12:43	254

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
		AP12	P1-LC-SP1PV1	P1_B3004, P1_LIT01	2020		
35	A402	AP07	P1-FC-SP1PV1	P1_B3005, P1_FT03		13:45	262
		AP25	P3-LC-CO2	P2_LCV01D			
36	A403	AP12	P1-LC-SP1PV1	P1_B3004, P1_LIT01		15:57	263
		AP25	P3-LC-CO2	P2_LCV01D			
37	A404	AP19	P2-SC-CO1PV1	P2_SCO, P2_SIT01		17:45	258
		AP14	P1-LC-CO1PV1	P1_LCV01D, P1_LIT01			
38	A405	AP20	P2-SC-SP1-ST	P2_AutoSD		20:47	120
		AP25	P3-LC-CO2	P2_LCV01D			
39	A501	AP03	P1-PC-CO1	P1_PCV01D	Jul. 30, 2020	11:16	172
		AP22	P2-TC-SP2	P2_VTR02			
40	A502	AP09	P1-FC-CO1PV1	P1_FCV03D, P1_FT03		13:30	258
		AP18	P2-SC-CO1	P2_SCO			
41	A503	AP12	P1-LC-SP1PV1	P1_B3004, P1_LIT01		16:05	256
		AP18	P2-SC-CO1	P2_SCO			
42	A504	AP08	P1-FC-CO1	P1_FCV03D		17:45	120
		AP25	P3-LC-CO2	P2_LCV01D			
43	A505	AP11	P1-LC-SP1	P1_B3004		18:38	203
		AP20	P2-SC-SP1-ST	P2_AutoSD			
44	A506	AP19	P2-SC-CO1PV1	P2_SCO, P2_SIT01		20:42	153
		AP25	P3-LC-CO2	P2_LCV01D			
45	A507	AP20	P2-SC-SP1-ST	P2_AutoSD		23:13	79
		AP21	P2-TC-SP1	P2_VTR01			
46	A508	AP10	P1-FC-CO1-ST	P1_FCV03D	Jul. 31, 2020	01:15	51
		AP15	P1-LC-CO1-ST	P1_LCV01D			
47	A509	AP01	P1-PC-SP1	P1_B2016		02:01	241
		AP03	P1-PC-CO1	P1_PCV01D			
48	A510	AP11	P1-LC-SP1	P1_B3004		09:54	262
		AP14	P1-LC-CO1PV1	P1_LCV01D, P1_LIT01			
49	A511	AP23	P2-TC-SP3	P2_RTR		10:40	120
		AP25	P3-LC-CO2	P2_LCV01D			
50	A512	AP06	P1-FC-SP1	P1_B3005		11:21	262
		AP09	P1-FC-CO1PV1	P1_FCV03D, P1_FT03			

HAI 20.07

HAI 20.07 includes two CSV files as training datasets and two CSV files as testing datasets. The time-series data in each CSV file satisfies time continuity and includes 63 columns. The first column represents the observed time in the “yyyy-MM-dd hh:mm:ss” format, and the remaining 59 columns provide the recorded SCADA data points. The last four columns provide data labels for whether an attack occurred or not. Out of these columns, the attack column is applicable to all processes and the other three columns are applicable to the corresponding control processes.

NORMAL OPERATION

The normal operations of the first training dataset (train1.csv) are given below, where all the SP change commands were delivered at the start of each day.

No	Setpoint						Start Time
	P1_B2004 (Pressure SP)	P1_B3004 (Level SP)	P1_B3005 (Flowrate SP)	P1_B4002 (Temperature SP)	P4_ST_PS (Scheduled Power)	P4_HT_PS (Scheduled Power)	
1	0.1 (±0.002)	460 (±20)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	7:00 (±0)
2	0.03 (±0.002)	400 (±8)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	9:00 (±0)
3	0.1 (±0.002)	400 (±8)	1,100 (±22)	31 (±1)	0 (±0)	0 (±0)	11:00 (±0)
4	0.1 (±0.002)	400 (±8)	1,000 (±100)	32 (±0)	0 (±0)	0 (±0)	13:00 (±0)
5	0.1 (±0.002)	400 (±8)	1,100 (±22)	32 (±0)	50 (±5)	0 (±0)	15:00 (±0)

The normal operations of the second training dataset (train2.csv) are given below.

No	Setpoint						Start Time
	P1_B2004 (Pressure SP)	P1_B3004 (Level SP)	P1_B3005 (Flowrate SP)	P1_B4002 (Temperature SP)	P4_ST_PS (Scheduled Power)	P4_HT_PS (Scheduled Power)	
1	0.03 (±0.002)	400 (±8)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	00:00 (±0)
2	0.1 (±0.002)	450 (±20)	1,100 (±22)	32 (±0)	0 (±0)	0 (±0)	10:00 (±0)
3	0.1 (±0.002)	400 (±8)	1,100 (±22)	32 (±1)	0 (±0)	0 (±0)	14:00 (±0)
4	0.1 (±0.002)	400 (±8)	1,000 (±100)	32 (±0)	0 (±0)	0 (±0)	16:00 (±0)
5	0.1 (±0.002)	400 (±8)	1,100 (±22)	32 (±0)	50 (±5)	0 (±0)	22:00 (±0)

ATTACK OPERATION

A total of 38 attacks were conducted, including 14 attack primitives and 14 combinations of attacks designed to simultaneously perform two attack primitives.

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
1	A101	AP12	P1-LC-SP1PV1	P1_B3004, P1_LIT01	Oct. 29, 2019	13:40	370
2	A102	AP13	P1-LC-CV1	P1_LCV01D		14:35	312

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
3	A103	AP14	P1-LC-CV1PV1	P1_LCV01D, P1_LIT01		15:45	868
4	A104	AP06	P1-FC-SP1	P1_B3005		16:30	262
5	A105	AP11	P1-LC-SP1	P1_B3004	Oct. 30, 2019	08:50	371
6	A106	AP01	P1-PC-SP1	P1_B2016		09:40	334
7	A107	AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01		10:35	504
8	A108	AP03	P1-PC-CV1	P1_PCV01D		11:37	268
9	A109	AP04	P1-PC-CV1PV1	P1_PCV01D, P1_PIT01		12:30	518
10	A110	AP17	P2-SC-SP1PV1	P2_SD01, P2_SIT01		14:30	370
11	A111	AP26	P3-LC-SP1CV1	P3_LH01, P3_LCP01		15:35	180
12	A112	AP27	P3-LC-SP2CV2	P3_LL01, P3_LCV01		16:33	154
13	A113	AP16	P2-SC-SP1	P2_SD01	Oct. 31, 2019	08:42	348
14	A114	AP17	P2-SC-SP1PV1	P2_SD01, P2_SIT01		10:30	518
		AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01			
15	A115	AP16	P2-SC-SP1	P2_SD01		11:33	346
		AP03	P1-PC-CV1	P1_PCV01D			
16	A116	AP17	P2-SC-SP1PV1	P2_SD01, P2_SIT01		13:25	368
17	A117	AP17	P2-SC-SP1PV1	P2_SD01, P2_SIT01		14:30	396
		AP14	P1-LC-CV1PV1	P1_LCV01D, P1_LIT01			
18	A118	AP16	P2-SC-SP1	P2_SD01		15:41	348
		AP06	P1-FC-SP1	P1_B3005			
19	A119	AP26	P3-LC-SP1CV1	P3_LH01, P3_LCP01		16:29	398
		AP01	P1-PC-SP1	P1_B2016			
20	A201	AP26	P3-LC-SP1CV1	P3_LH01, P3_LCP01	Nov. 1, 2019	09:29	560
		AP12	P1-LC-SP1PV1	P1_B3004, P1_LIT01			
21	A202	AP26	P3-LC-SP1CV1	P3_LH01, P3_LCP01		10:41	310
		AP13	P1-LC-CV1	P1_LCV01D			
22	A203	AP26	P3-LC-SP1CV1	P3_LH01, P3_LCP01		11:23	180
23	A204	AP11	P1-LC-SP1	P1_B3004		12:31	506

No	ID	Attack			Start Time		Duration (sec)
		Scenario	Target Controller	Target Point(s)			
		AP07	P1-FC-SP1PV1	P1_B3005, P1_FT03			
24	A205	AP03	P1-PC-CV1	P1_PCV01D		13:41	580
		AP07	P1-FC-SP1PV1	P1_B3005, P1_FT03			
25	A206	AP01	P1-PC-SP1	P1_B2016		14:23	310
26	A207	AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01		15:31	520
		AP06	P1-FC-SP1	P1_B3005			
27	A208	AP07	P1-FC-SP1PV1	P1_B3005, P1_FT03		16:18	560
28	A209	AP27	P3-LC-SP2CV2	P3_LL01, P3_LCV01		17:20	520
		AP02	P1-PC-SP1PV1	P1_B2016, P1_PIT01			
29	A210	AP01	P1-PC-SP1	P1_B2016	Nov. 4, 2019	15:31	410
		AP06	P1-FC-SP1	P1_B3005			
30	A211	AP24	P3-LC-SP2CV2	P3_SP02, P3_LCV01		17:20	520
		AP01	P1-PC-SP1	P1_B2016			
31	A212	AP24	P3-LC-SP2CV2	P3_SP02, P3_LCV01	Nov. 5, 2019	09:30	380
		AP13	P1-LC-CV1	P1_LCV01D			
32	A213	AP24	P3-LC-SP2CV2	P3_SP02, P3_LCV01		10:20	290
		AP06	P1-FC-SP1	P1_B3005			
33	A214	AP16	P2-SC-SP1	P2_SD01		11:23	340
34	A215	AP16	P2-SC-SP1	P2_SD01		12:30	340
		AP27	P3-LC-SP2CV2	P3_LL01, P3_LCV01			
35	A216	AP16	P2-SC-SP1	P2_SD01		14:45	2,880
		AP11	P1-LC-SP1	P1_B3004			
36	A217	AP11	P1-LC-SP1	P1_B3004		16:20	330
		AP01	P1-PC-SP1	P1_B2016			
37	A218	AP13	P1-LC-CV1	P1_LCV01D		17:23	310
38	A219	AP13	P1-LC-CV1	P1_LCV01D	Nov. 6, 2019	08:58	310
		AP03	P1-PC-CV1	P1_PCV01D			

CASE STUDIES

We provide Python NetworkX graph data (JSON format) for boiler system along with the release of a HAI 23.05 and HAIEnd 23.05. You can create the separate graphs for the two subsystems: the digital subsystem and physical subsystem and also merge them by connecting the sensor and actuator nodes of the two subsystems. The NetworkX graphs helps in analyzing and optimizing anomaly detection performance. Several case studies with this tool are as follows.

Data flow graph for digital subsystem

PCL-based data flow graphs are suitable for user purposes using typical graph analysis methodologies. For example, when analyzing attack scenario initiated by the digital subsystem (DCS, HMI, EWS and so on), the reachability to all traversal results from the start node (attack initiation) to the end node (attack target) can be used, and extracts and selects all reachable paths as traversal results. Furthermore, an attack propagation chain (APC), a path composed of nodes and edges affected by the attack, is provided for sophisticated analysis for anomaly detector.

CASE I: ATTACK PROPAGATION CHAIN

An APC describes a path composed of nodes and edges being attacked and applies forward analysis to the graph to determine the propagation path in the forward direction with the analysis target node as the starting point. It then determines which nodes and edges are situated along the propagation path affected by the attack. APCs can also be used to identify hidden targets and scopes when constructing an attack scenario.

When HMI is compromised by an attacker, the impact can be identified using an APC as indicated by the bold path as shown in Figure 16. Attack propagation chain is derived as two distinct paths. The first path (In order of path: 1001.7, 1001.8, 1001.20, 1001.21, 1001.22/1001.23, and 1001.24) refers to a situation in which an attacker spreads the influence through keyboard manipulation of the HMI. The second path (In order of path: 1001.5, 1001.14, 1001.15, 1001.16, 1001.17, 1001.18, and 1001.19) indicates the effect of manipulation the setpoint. In particular, here is an APC for one of the attack scenarios described in previous section. When manipulating internal point at DM-PCV01-D in 1001.21 (i.e., ATTACK SCENARIO AE05), the impact can be identified using an APC as indicated by the bold path as shown in Figure 17. Once the attack proceeds at entry point, related data (control and measurement values) of PCV01 and PCV02 are affected as well.

CASE II: ROOT-CAUSE ANALYSIS

In contrast to APC, backward analysis of a graph can elucidate the root-cause of an attack. The attack entry point can be identified from the node corresponding to the path end by tracing the path backward from a specific node. For example, when backward analysis is applied to the graph for anomaly of PIT01 as shown in Figure 18, root-cause (HMI, HIL, EWS) can be provided.

Physics-related flow graph for physical subsystem

A Physics-related flow graph is a visual representation of the interconnected components and processes within a physical subsystem. It provides a clear and organized overview of how energy and information flow through different elements of the system, enabling the analysis and understanding of its behavior.

In the context of a water-based heating system, such as a boiler, the graph would depict the flow of energy through hydrodynamics and thermodynamics. Nodes in the graph represent specific components within the boiler system, while edges represent the transfer of energy or interactions between these nodes. In this case, the energy transfer within the boiler is facilitated by the principle of hydrodynamics and thermodynamics. Hydrodynamics would represent the movement of water and the energy transfer associated with fluid flow. As show in Fig. 19, this could include the flow of water through pipes, valves, and other components within the boiler systems. Thermodynamics, on the other hand, would encompass processes such as heat transfer, energy conversion, and the overall efficiency of the heating system. This would involve interactions between the water, heat source, and various components like heat exchangers and pumps.

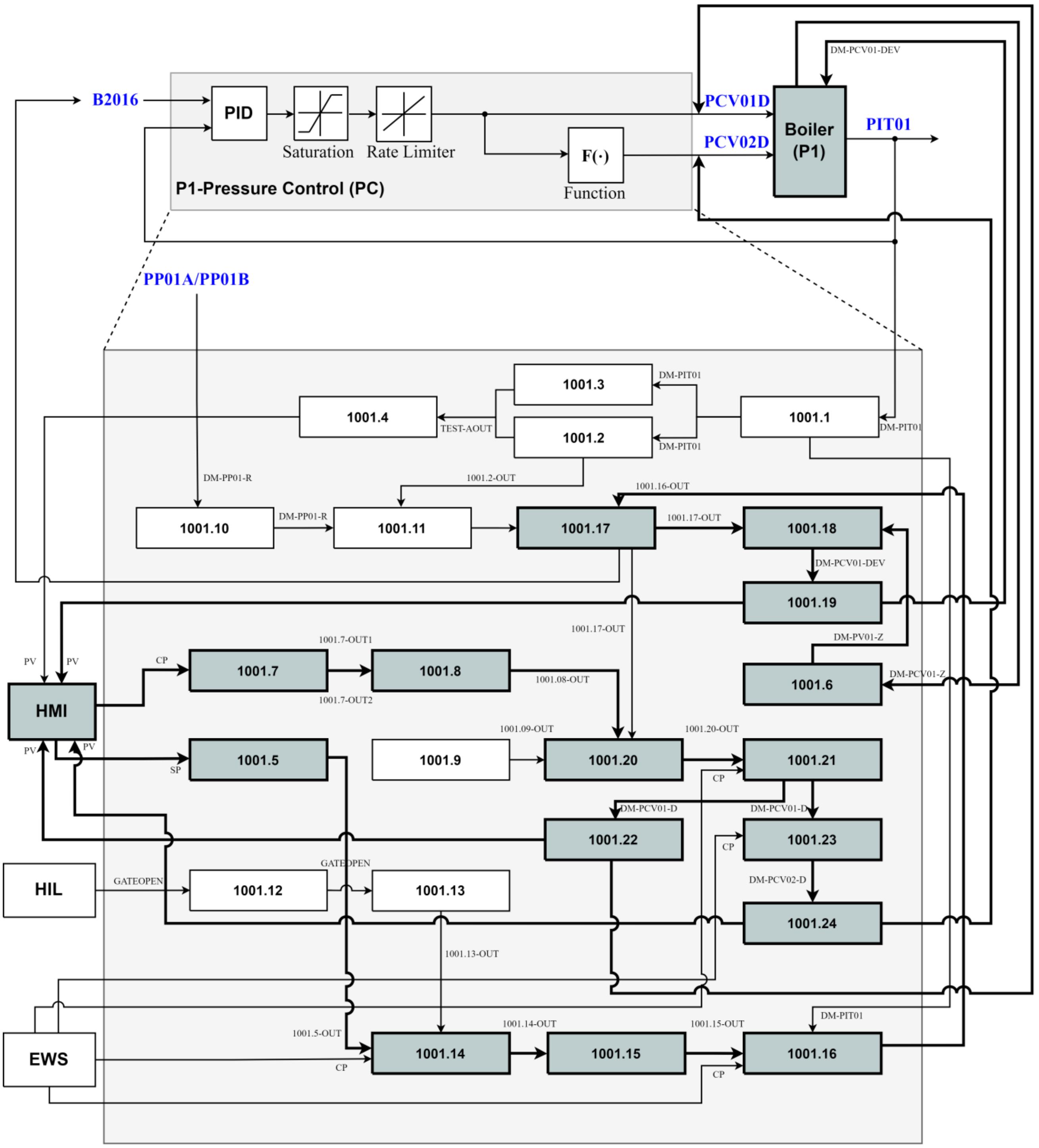

FIGURE 16. ATTACK PROPAGATION CHAIN WHEN COMPROMISED HMI

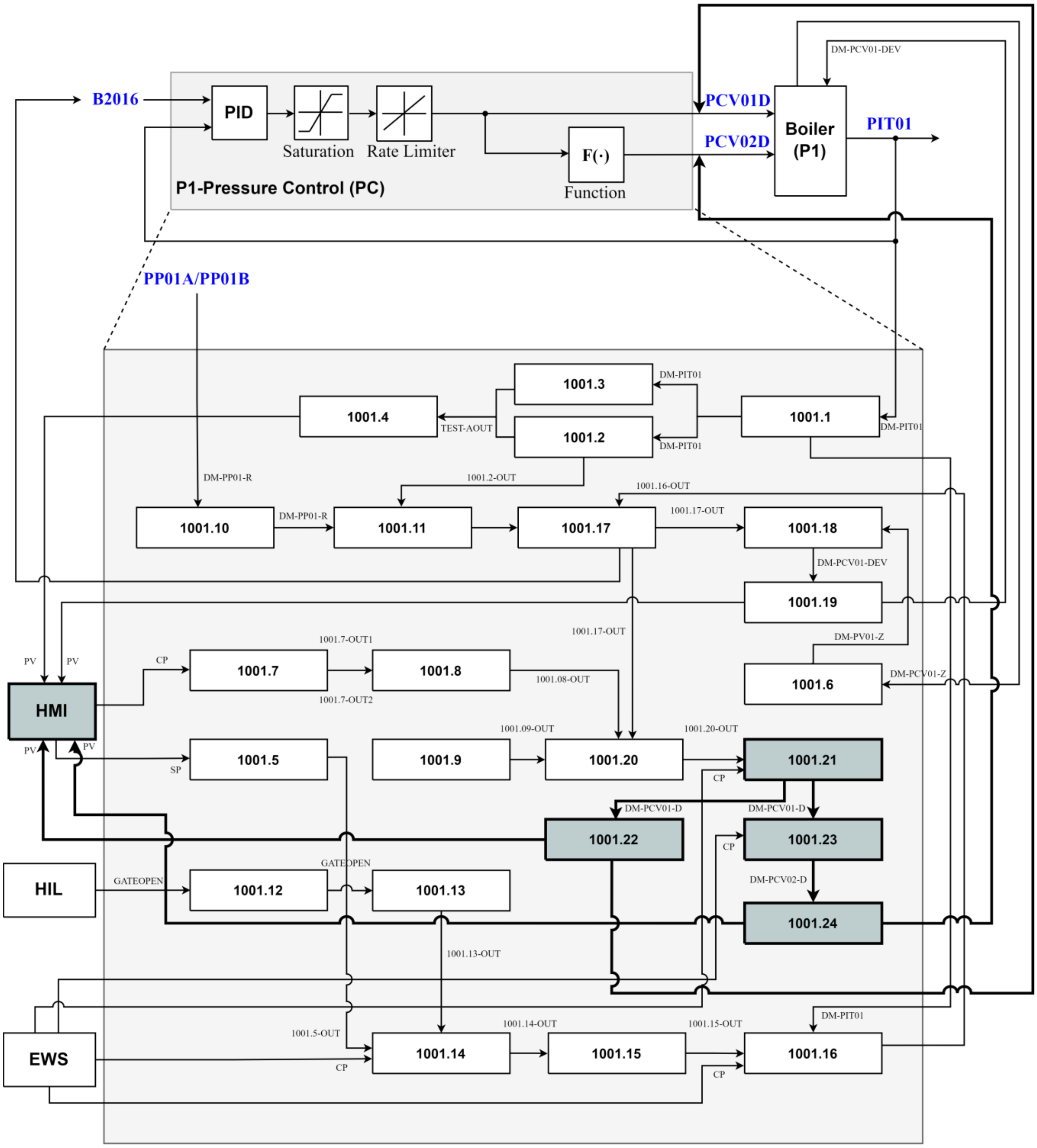

FIGURE 17. ATTACK PROPAGATION CHAIN WHEN INJECTING INTERNAL POINT ATTACK AT DM-PCV01-D IN 1001.21 (ATTACK SCENARIO AE05)

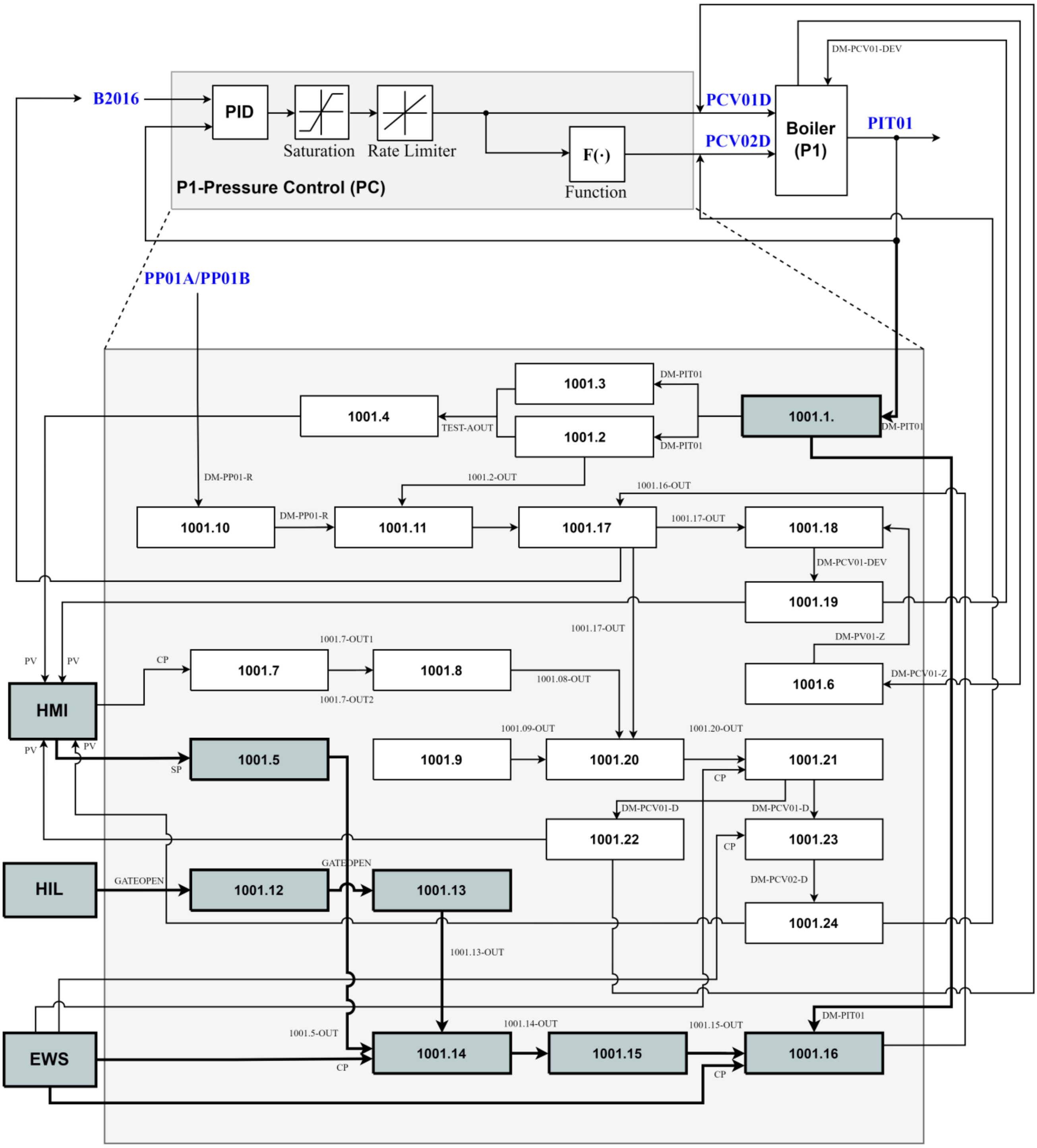

FIGURE 18. ROOT-CAUSE ANALYSIS WHEN DETECTING AN ANOMALY AT PIT01

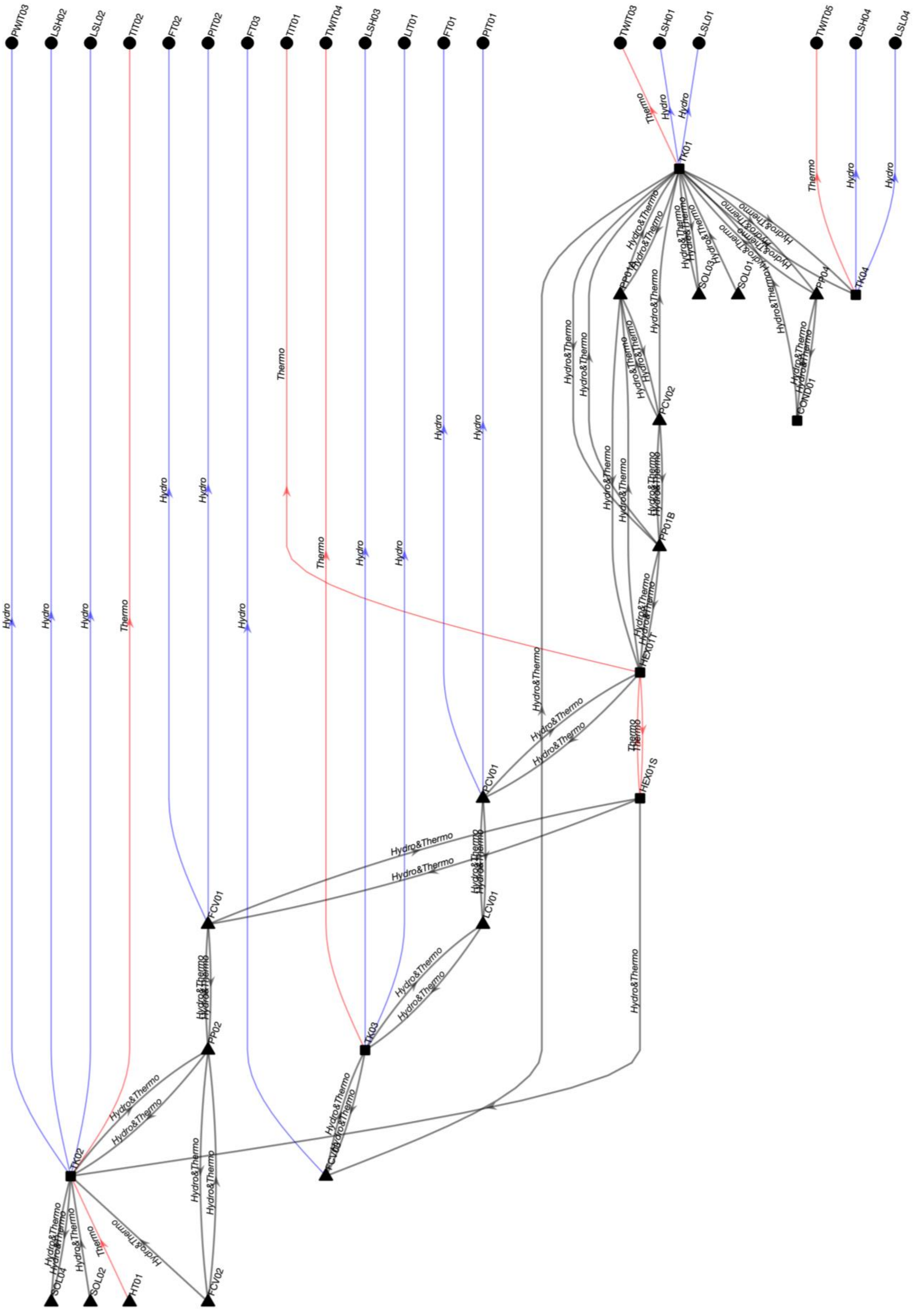

FIGURE 19. PHYSICS-REALATED FLOW GRAPH FOR BOILER’S PHYSCAL SUBSYSTEM

CITATION

Please cite the sources below if you are referencing any of the HAI datasets, performance matric, and competitions. Please do not hesitate to share your results with us.

Datasets

[HAI 22.04] Hyeok-Ki Shin, Woomyo Lee, Jeong-Han Yun and Byung-Gil Min, "ICS security dataset", 2022. *Available at:*

- *GitHub: https://github.com/icsdataset/hai*
- *Kaggle: https://kaggle.com/icsdataset/hai-security-dataset*

[HAI 21.03] Hyeok-Ki Shin, Woomyo Lee, Jeong-Han Yun and Byung-Gil Min, "Two ICS Security Datasets and Anomaly Detection Contest on the HIL-based Augmented ICS Testbed", *In Proceedings of Cyber Security Experimentation and Test (CSET `21), Association for Computing Machinery, pp 36-40, 2021.*

[HAI 20.07] Hyeok-Ki Shin, Woomyo Lee, Jeong-Han Yun and Hyoungchun Kim, "HAI 1.0: HIL-based Augmented ICS Security Dataset," *In Proceedings of the 13th USENIX Conference on Cyber Security Experimentation and Test (CSET `20), USENIX Association, 2020.*

Performance Analysis

[eTaPR] Won-Seok Hwang, Jeong-Han Yun, Jonguk Kim, and Byung Gil Min, "Do you know existing accuracy metrics overate time-series anomaly detection?", *In Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing (SAC`22),* Association for Computing Machinery, pp403-412, 2022.

- *GitHub:* https://github.com/saurf4ng/eTapR

[DFG] Seungoh Choi, Hyeok-Ki Shin, Woomyo Lee, Jeong-Han Yun and Byung-Gil Min, "Dataflow-based Control Process Identification for ICS Dataset Development", *In Proceedings of the 15th Workshop on Cyber Security Experimentation and Test (CSET `22)*, Association for Computing Machinery, pp.54-58, 2022.

Competitions/Baseline

[HAICon] We held an AI contest, namely HAICon, to revitalized research, discover ideas, and improve HAI dataset more. You can find the winner's codes and baseline codes on the official website below.

- HAICon 2021: https://dacon.io/en/competitions/official/235757/codeshare
- HAICon 2020: https://dacon.io/en/competitions/official/235624/codeshare

ABBREVIATIONS

A

APC	ATTACK PROPAGATION CHAIN

C

CV	CONTROL VARIABLE
CC	COOLING CONTROLLER
CO	CONTROL OUTPUT
CP	CONTROL PARAMETERER

D

DCS	DISTRIBUTED CONTROL SYSTEM

F

FC	FLOW CONTROLLER
FCV	FLOW CONTROL VALVE
FIT	FLOW INDICATOR TRANSMITTER
FT	FLOW TRANSMITTER

H

HH	HIGH HIGH
HIL	HARDWARE-IN-THE-LOOP
HMI	HUMAN MACHINE INTERFACE

L

LC	LEVEL CONTROLLER
LCV	LEVEL CONTROL VALVE
LIT	LEVEL INDICATOR TRANSMITTER
LL	LOW LOW
LLH	LIQUID LEVEL [HIGH]
LLL	LIQUID LEVEL [LOW]
LLN	LIQUID LEVEL [NORMAL]
LSH	LEVEL SWITCH [HIGH]
LSHL	EVEL SWITCH [HIGH/LOW]
LSL	LEVEL SWITCH [LOW]
LT	EVEL TRANSMITTER

P

PC	PRESSURE CONTROLLER
PCL	PROCESS CONTROL LOOP
PCV	PRESSURE CONTROL VALVE
PIT	PRESSURE INDICATOR TRANSMITTER

PLC	PROGRAMMABLE LOGIC CONTROLLER
PV	PROCESS VARIABLE

S

SC	SPEED CONTROLLER
SI	SPEED INDICATOR
SIT	SPEED-INDICATOR TRANSMITTER
SP	SETPOINT
SS	STEAM SUPPLY

T

TCV	TEMPERATURE CONTROL VALVE
TIT	TEMPERATURE-INDICATOR TRANSMITTER
TT	TEMPERATURE TRANSMITTER

V

VT	VIBRATION TRANSMITTER