



Industrial Cybersecurity

Opportunities and challenges
in Digital Transformation



ANDREY LAVRENTYEV


Kaspersky Lab
Russia

- Head of Technology Research Department, Future Technologies
- Runs cyber-security researches, heavily based on AI
- Created MLAD system that detects attacks and faults in industrial processes
- Recent successful implementation on TANECO oil refinery plant

[linkedin.com/in/andrey-lavrentyev-2907b14b](https://www.linkedin.com/in/andrey-lavrentyev-2907b14b)



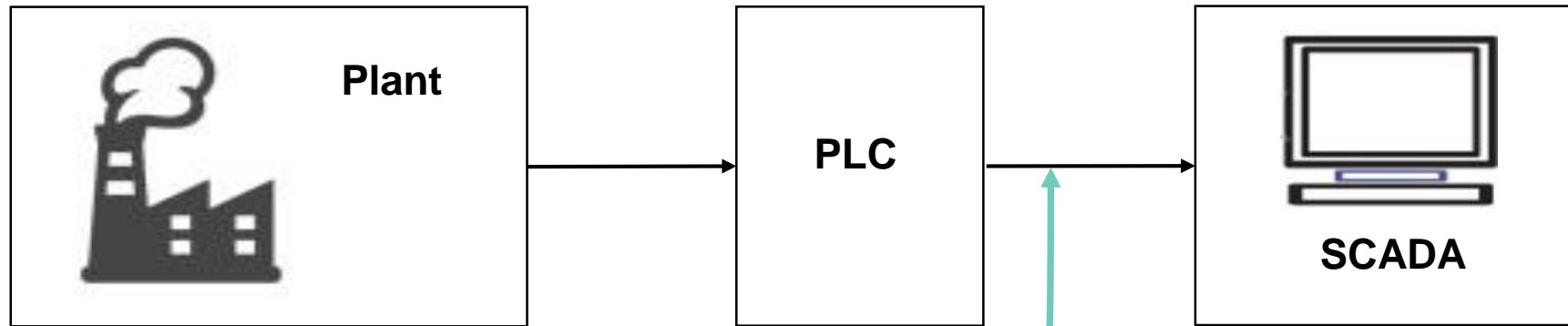
ICS Operational Technology Protection with Machine Learning

The background of the slide is a dark, high-tech industrial environment. It features several large, vertical cylindrical tanks or reactors. These tanks are illuminated with vibrant green and orange light, suggesting active processes or data flow. The surrounding area is filled with a complex network of pipes, walkways, and structural elements, all rendered in a dark, metallic style. The overall aesthetic is that of a modern, data-driven industrial plant.

Andrey Lavrentyev

Head of Technology Research Department,
Future Technologies,
Kaspersky Lab

ICS is a Cyber-Physical System



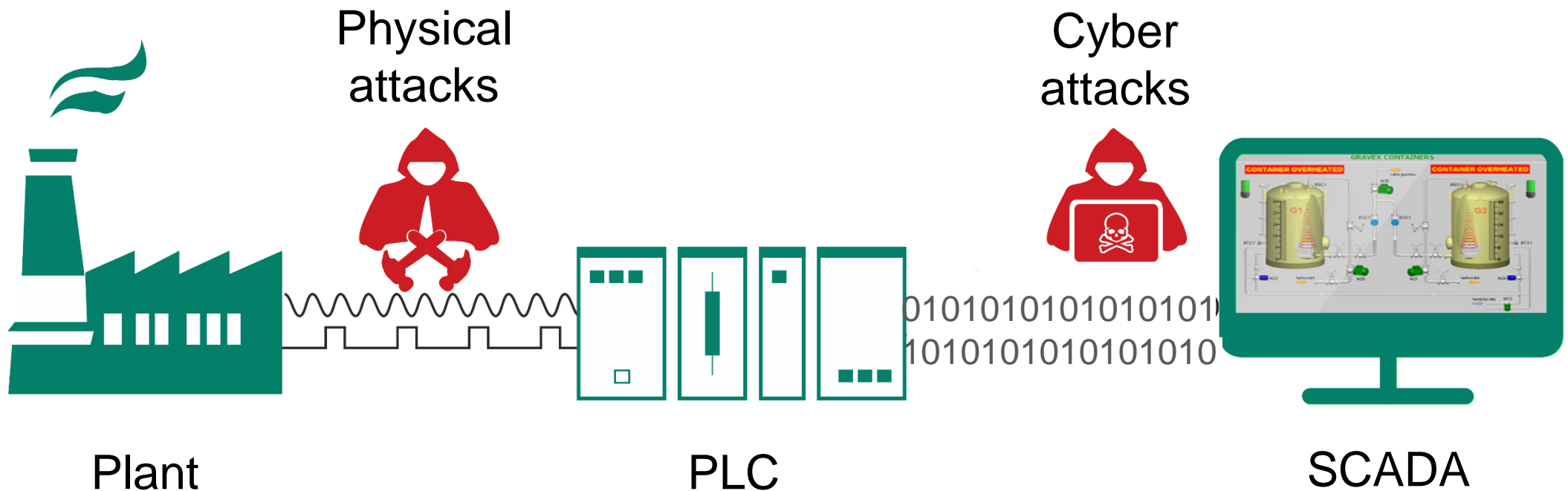
Typical industrial telemetry data:

- Multi-variate $\sim 10^4$ channels
- Real-time updates ~ 1 sec
- Big history \sim years
- Noise, time-jitter, gaps, faults

ICS under Attack

Attacks may **target**:

- information technology (**IT**) or
- operational technology (**OT**)

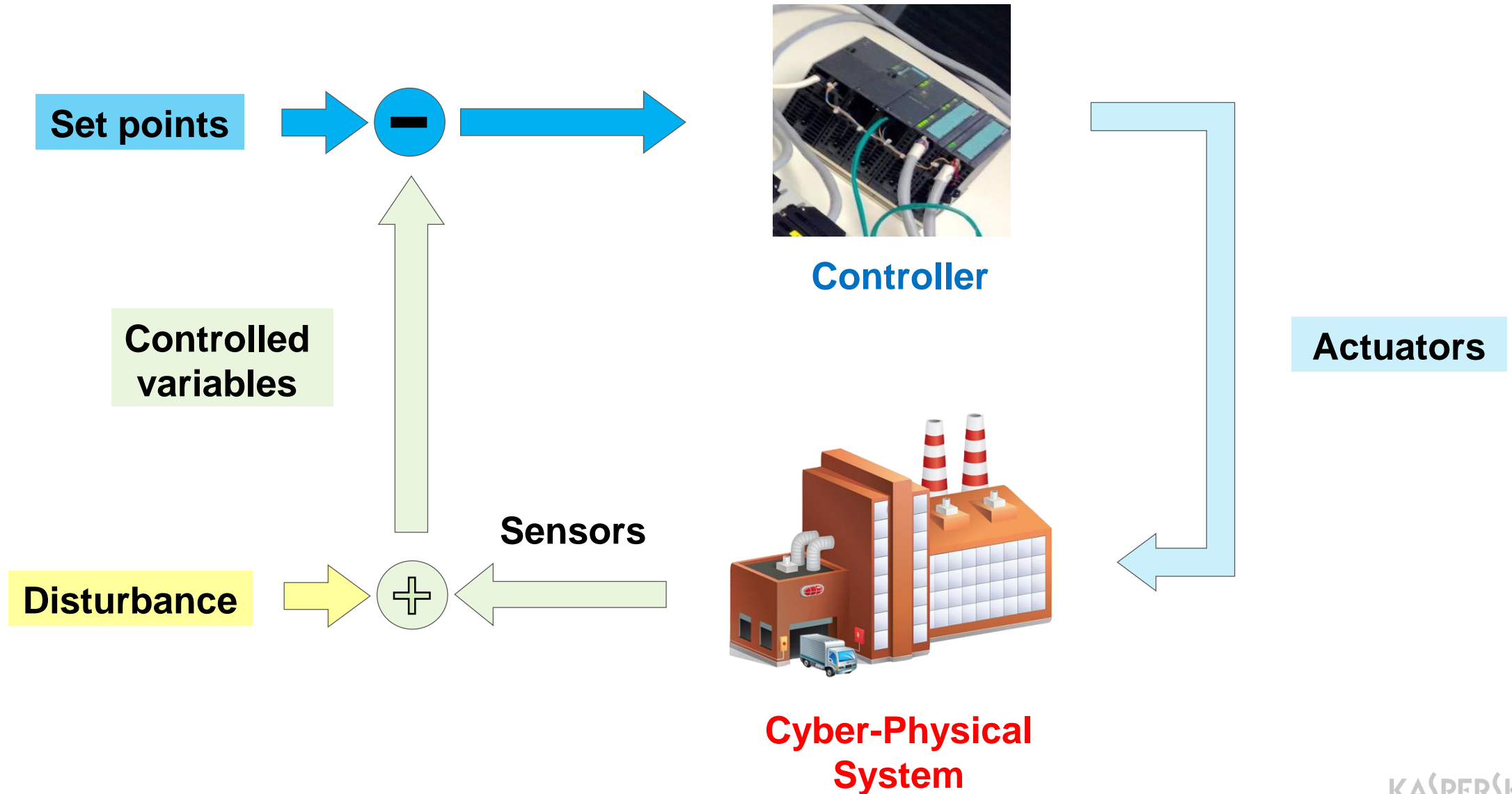


Attacks on OT are the most dangerous

- threat to a human life
- quick damage to physical equipment
- severe financial losses



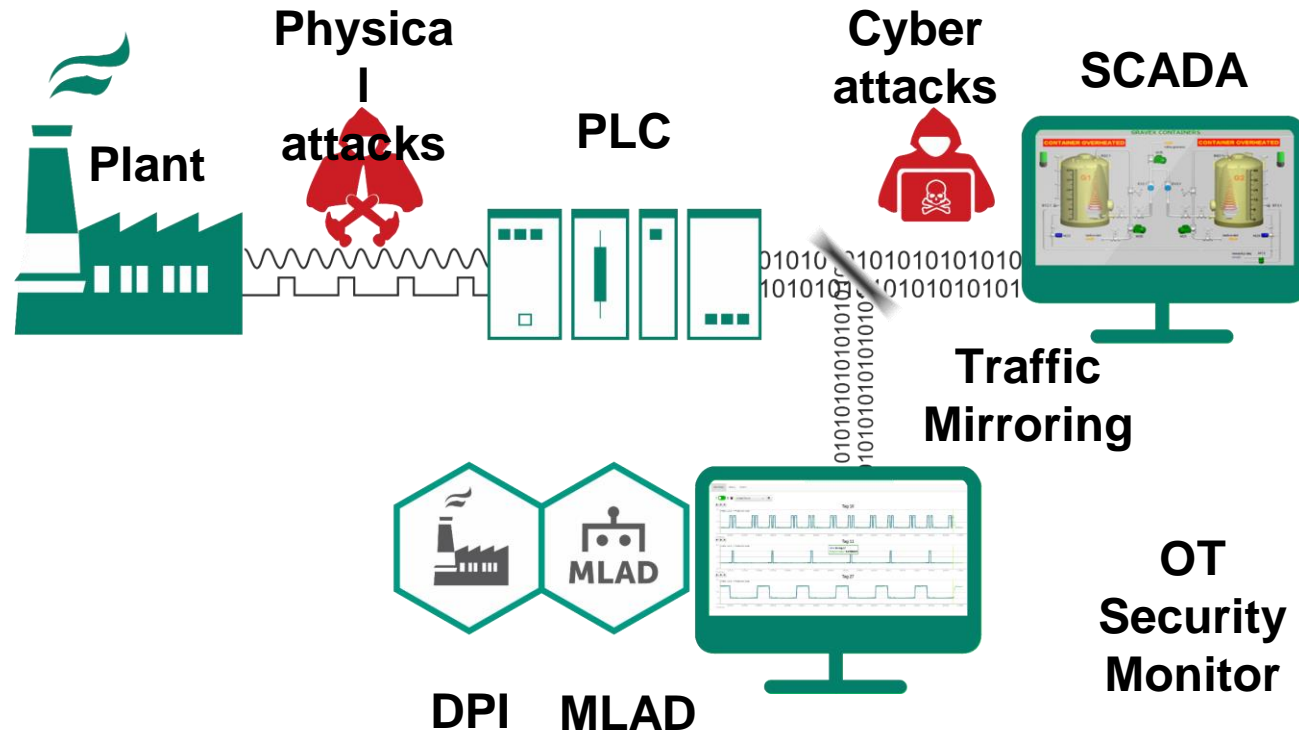
Control Loop for the Cyber-Physical System



How to detect attacks on OT?

1. Industrial **signals** (sensor and actuator values, control logic parameters **are correlated and governed by physical laws and control logic**)
2. Any (intentional or unintentional) modification in **one** signal causes **changes to correlated** with it signals
3. These correlations in signals can be **learned** and changes **detected with ML**

MLAD - Machine Learning for Anomaly Detection



MLAD:

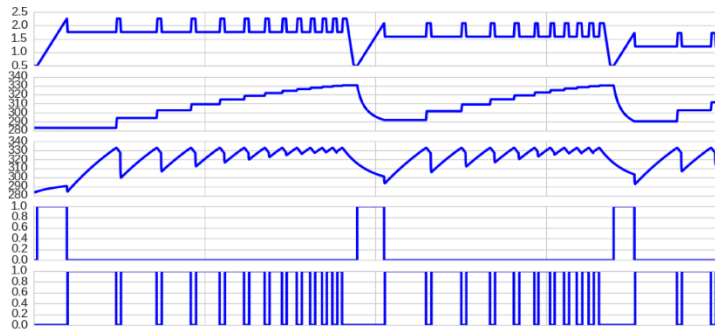
- ✓ Непрерывно строит прогнозные значения телеметрии процессов
- ✓ Непрерывно сравнивает прогноз с реальными значениями
- ✓ Непрерывно осуществляет мониторинг ошибки прогноза
- ✓ Автоматически интерпретирует найденные аномалии
- ✓ Имеет возможность дообучаться на новых данных

Neural Network as a Forecasting Function

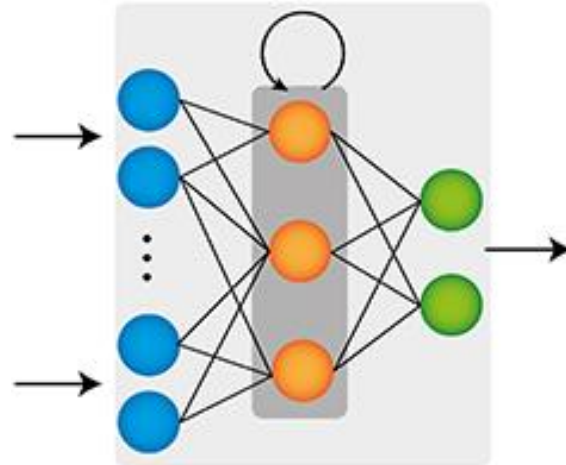
$$f : \mathbb{R}^{L \times m} \rightarrow \mathbb{R}^{\tilde{L} \times m}$$

Input: Multivariate Time Series

$$\mathbf{x}_t = (x_{t1}, \dots, x_{ti}, \dots, x_{tm})$$



Neural Network



Prediction, Observation and Error



Input window

$$W_k = (\mathbf{x}_t, \mathbf{x}_{t+1}, \dots, \mathbf{x}_{t+L})$$

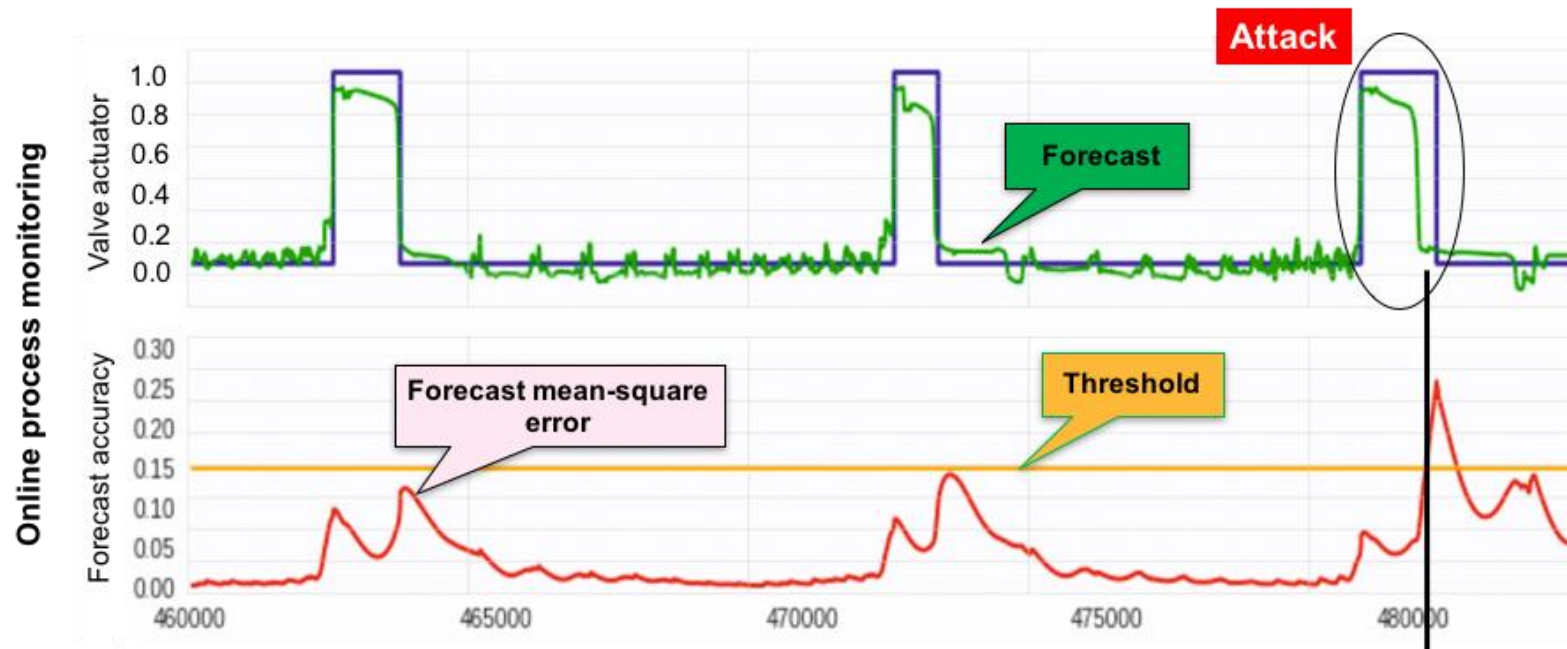
Forecast window

$$\tilde{W}_k = (\tilde{\mathbf{x}}_{\tilde{t}}, \tilde{\mathbf{x}}_{\tilde{t}+1}, \dots, \tilde{\mathbf{x}}_{\tilde{t}+\tilde{L}})$$

Anomaly Detection (AD)

$$\text{Error: } = \frac{1}{m} \sum_{i=1}^m w_i |\tilde{x}_{ti} - x_{ti}|^p$$

Error > **Threshold*** → **Anomaly**



* Threshold can be defined as 99.99 percentile on training dataset under normal operation conditions

Anomaly Interpretation

$$\text{Error:} = \frac{1}{m} \sum_{i=1}^m w_i |\tilde{x}_{ti} - x_{ti}|^p$$

Sort all variables against individual errors :

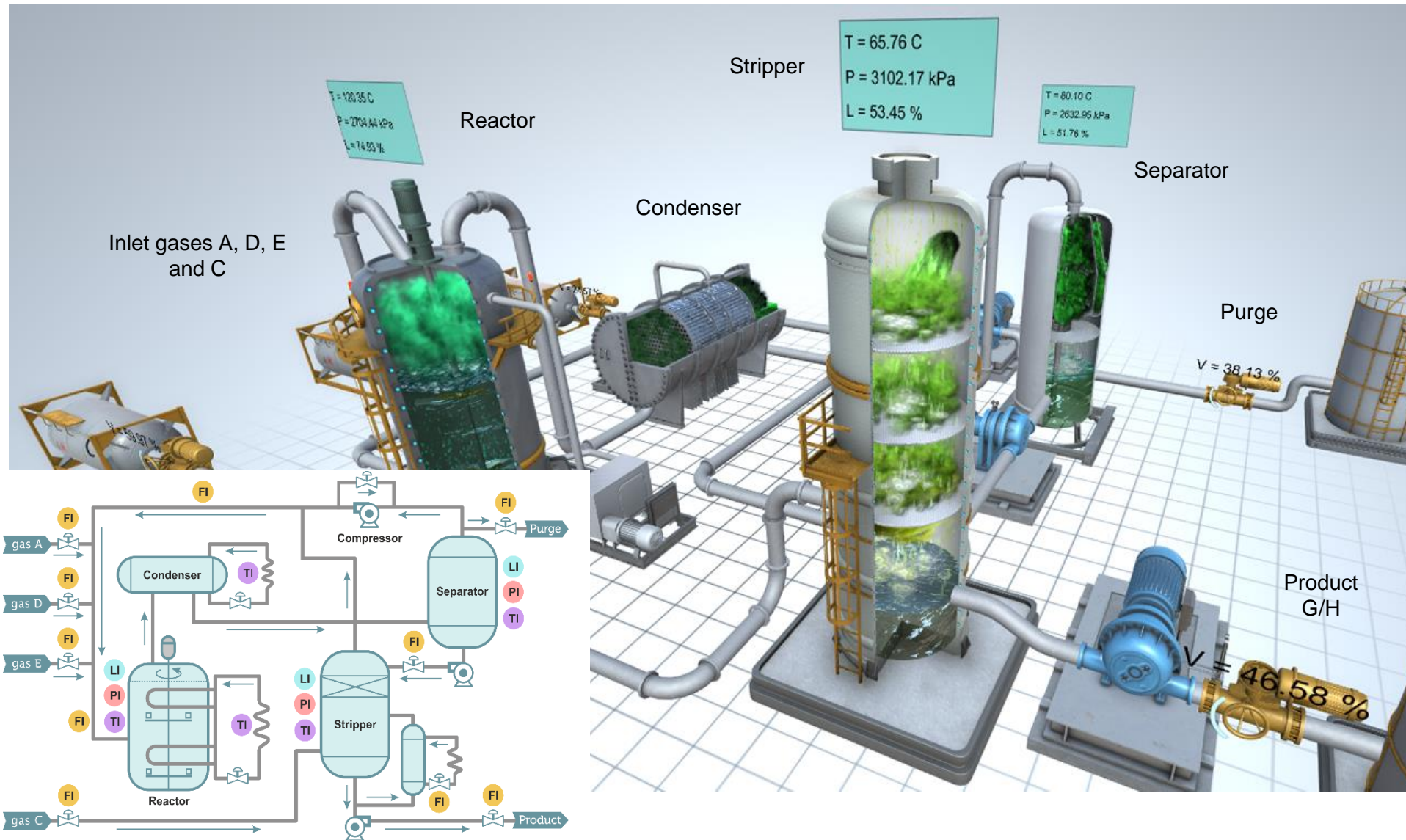
- determine most impacted variables
- diagnose root of anomaly*

$$w_i |\tilde{x}_{ti} - x_{ti}|^p$$

* Impact may be caused either by cyber or physical attack or by equipment fault.

Examples:

1. Chemical plant: Tennessee Eastman Process (TEP)

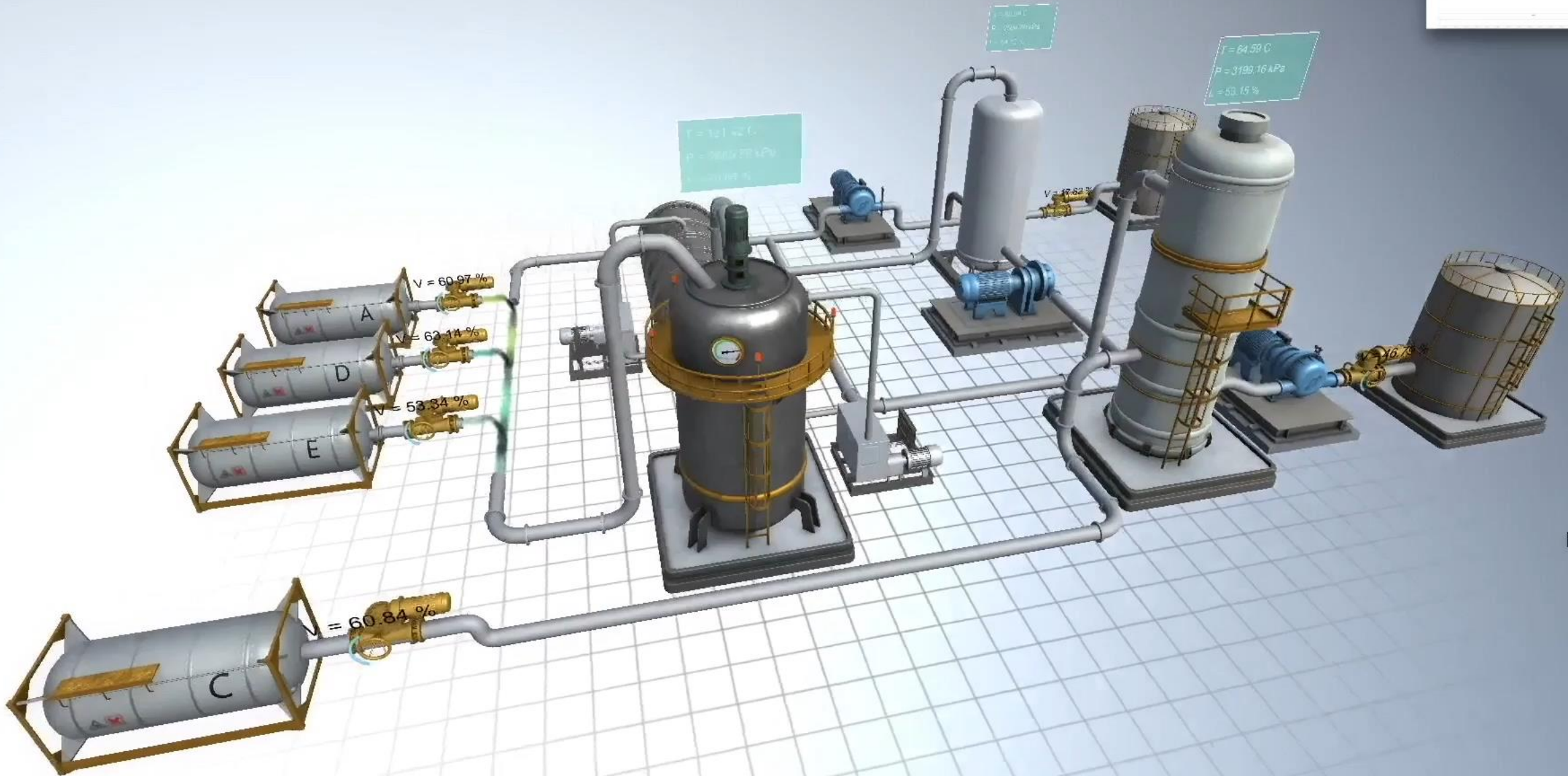


Tennessee Eastman Process
Cyber-Attack Detection
with ML

Industry
Cyber
Safegu
Progre



Stream	Phase	Temperature (C)	Pressure (MPa)	Flow Rate (kg/h)
1	Gas	121.47	3005.77	10.97
2	Gas	121.47	3005.77	10.97
3	Gas	121.47	3005.77	10.97
4	Gas	121.47	3005.77	10.97
5	Gas	121.47	3005.77	10.97
6	Gas	121.47	3005.77	10.97
7	Gas	121.47	3005.77	10.97
8	Gas	121.47	3005.77	10.97
9	Gas	121.47	3005.77	10.97
10	Gas	121.47	3005.77	10.97
11	Gas	121.47	3005.77	10.97
12	Gas	121.47	3005.77	10.97
13	Gas	121.47	3005.77	10.97
14	Gas	121.47	3005.77	10.97
15	Gas	121.47	3005.77	10.97
16	Gas	121.47	3005.77	10.97
17	Gas	121.47	3005.77	10.97
18	Gas	121.47	3005.77	10.97
19	Gas	121.47	3005.77	10.97
20	Gas	121.47	3005.77	10.97
21	Gas	121.47	3005.77	10.97
22	Gas	121.47	3005.77	10.97
23	Gas	121.47	3005.77	10.97
24	Gas	121.47	3005.77	10.97
25	Gas	121.47	3005.77	10.97
26	Gas	121.47	3005.77	10.97
27	Gas	121.47	3005.77	10.97
28	Gas	121.47	3005.77	10.97
29	Gas	121.47	3005.77	10.97
30	Gas	121.47	3005.77	10.97
31	Gas	121.47	3005.77	10.97
32	Gas	121.47	3005.77	10.97
33	Gas	121.47	3005.77	10.97
34	Gas	121.47	3005.77	10.97
35	Gas	121.47	3005.77	10.97
36	Gas	121.47	3005.77	10.97
37	Gas	121.47	3005.77	10.97
38	Gas	121.47	3005.77	10.97
39	Gas	121.47	3005.77	10.97
40	Gas	121.47	3005.77	10.97
41	Gas	121.47	3005.77	10.97
42	Gas	121.47	3005.77	10.97
43	Gas	121.47	3005.77	10.97
44	Gas	121.47	3005.77	10.97
45	Gas	121.47	3005.77	10.97
46	Gas	121.47	3005.77	10.97
47	Gas	121.47	3005.77	10.97
48	Gas	121.47	3005.77	10.97
49	Gas	121.47	3005.77	10.97
50	Gas	121.47	3005.77	10.97
51	Gas	121.47	3005.77	10.97
52	Gas	121.47	3005.77	10.97
53	Gas	121.47	3005.77	10.97
54	Gas	121.47	3005.77	10.97
55	Gas	121.47	3005.77	10.97
56	Gas	121.47	3005.77	10.97
57	Gas	121.47	3005.77	10.97
58	Gas	121.47	3005.77	10.97
59	Gas	121.47	3005.77	10.97
60	Gas	121.47	3005.77	10.97
61	Gas	121.47	3005.77	10.97
62	Gas	121.47	3005.77	10.97
63	Gas	121.47	3005.77	10.97
64	Gas	121.47	3005.77	10.97
65	Gas	121.47	3005.77	10.97
66	Gas	121.47	3005.77	10.97
67	Gas	121.47	3005.77	10.97
68	Gas	121.47	3005.77	10.97
69	Gas	121.47	3005.77	10.97
70	Gas	121.47	3005.77	10.97
71	Gas	121.47	3005.77	10.97
72	Gas	121.47	3005.77	10.97
73	Gas	121.47	3005.77	10.97
74	Gas	121.47	3005.77	10.97
75	Gas	121.47	3005.77	10.97
76	Gas	121.47	3005.77	10.97
77	Gas	121.47	3005.77	10.97
78	Gas	121.47	3005.77	10.97
79	Gas	121.47	3005.77	10.97
80	Gas	121.47	3005.77	10.97
81	Gas	121.47	3005.77	10.97
82	Gas	121.47	3005.77	10.97
83	Gas	121.47	3005.77	10.97
84	Gas	121.47	3005.77	10.97
85	Gas	121.47	3005.77	10.97
86	Gas	121.47	3005.77	10.97
87	Gas	121.47	3005.77	10.97
88	Gas	121.47	3005.77	10.97
89	Gas	121.47	3005.77	10.97
90	Gas	121.47	3005.77	10.97
91	Gas	121.47	3005.77	10.97
92	Gas	121.47	3005.77	10.97
93	Gas	121.47	3005.77	10.97
94	Gas	121.47	3005.77	10.97
95	Gas	121.47	3005.77	10.97
96	Gas	121.47	3005.77	10.97
97	Gas	121.47	3005.77	10.97
98	Gas	121.47	3005.77	10.97
99	Gas	121.47	3005.77	10.97
100	Gas	121.47	3005.77	10.97



2. Secure Water Treatment System (SWaT)



SWaT – testbed at Singapore University of Tech and Design

<https://itrust.sutd.edu.sg/research/testbeds/secure-water-treatment-swat/>

<https://itrust.sutd.edu.sg/dataset/>

SWaT processes:

P1: RAW water Supply and storage

P2: Pre-treatment

P3: Ultrafiltration and backwash

P4: De-Chlorination System

P5: Reverse Osmosis (RO)

P6: RO Permeate Transfer, UF Backwash and Cleaning

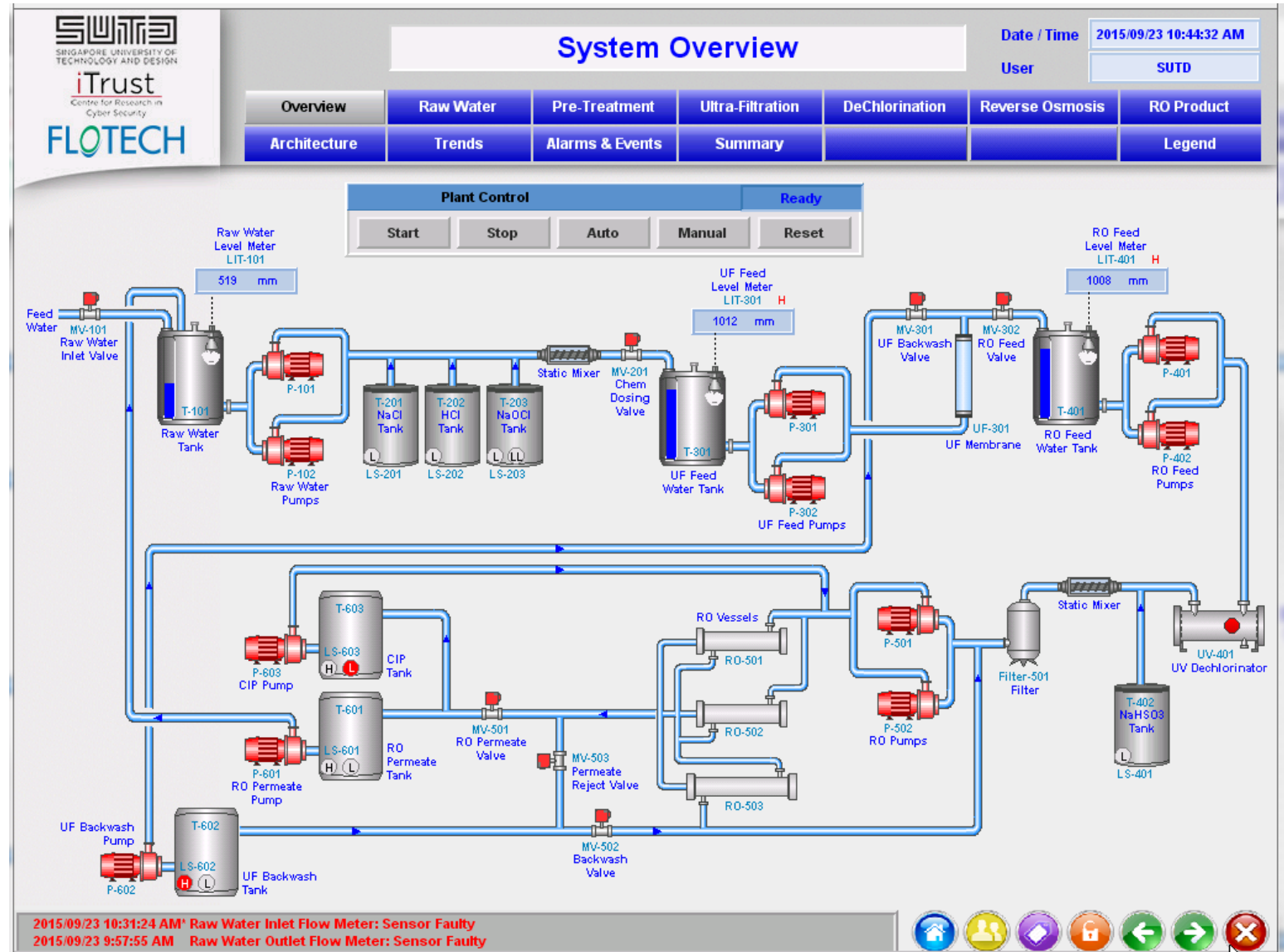
SWaT dataset:

25 sensors

26 actuators

7 days normal operation conditions

6 days 34 attacks with physical imp.



[Train model](#)

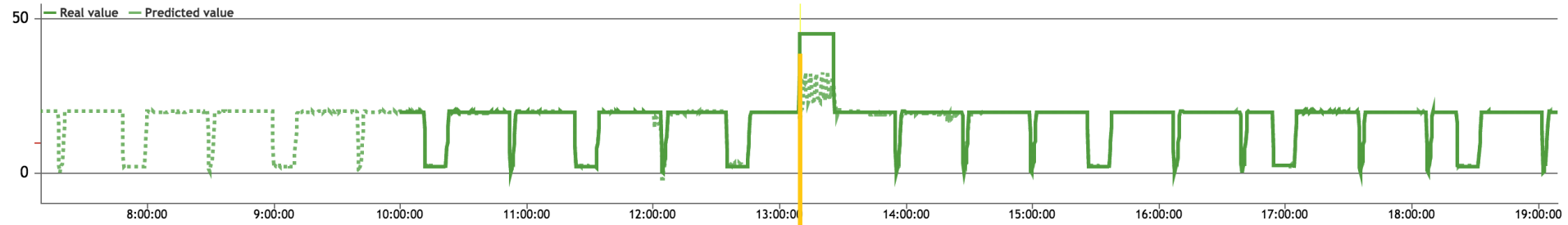
Name	Status	Progress	Switcher
+ SWaT dense model	Ready		Off <input checked="" type="checkbox"/> On ✕ ⚙️



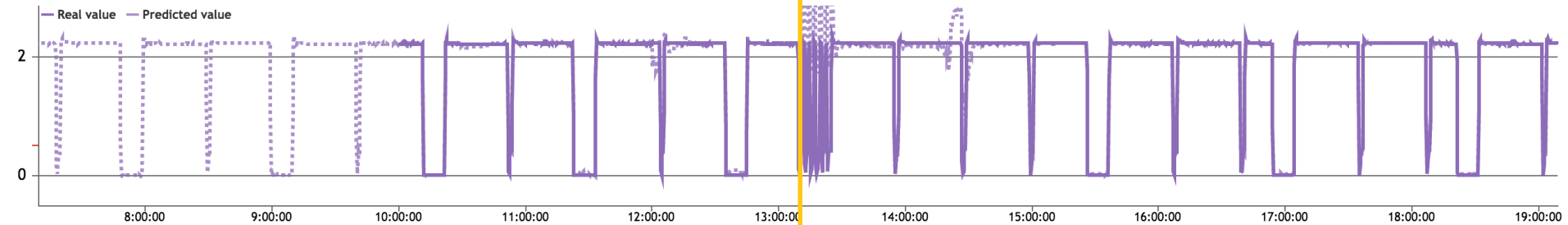
Detected Attack

← 12 hours → MSE ★ Sorted Tags Date: 2015-12-28

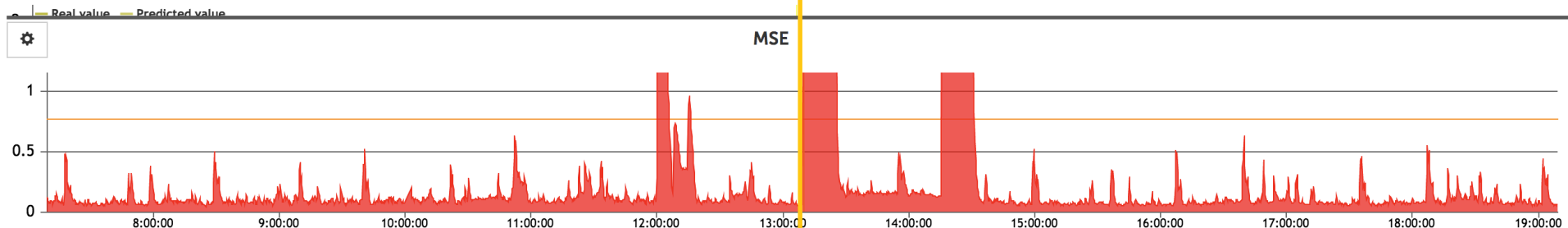
DPIT301



FIT301



PIT502



Why some of the attacks may be ignored?

$$\mathbf{Error} = \frac{1}{m} \sum_{i=1}^m w_i |\tilde{x}_{ti} - x_{ti}|^p$$

If attacked variable has **small impact** on other variables then only one individual error $w_i |\tilde{x}_{ti} - x_{ti}|^p$ will be quite large and summarized error may **not reach detection threshold**.

Attack with small impact



Detected attack:
23 out of 34

Not detected:
9 with small impacts

Correct interpretation:
22 out of 23 !!!

False positive:
3 as attack continuation

New anomalies:
7 anomalies

26	Single Stage Single Point Attacks
4	Single Stage Multi Point Attacks
2	Multi Stage Single Point Attacks
4	Multi Stage Multi Point Attacks

Attack #	Start Time	End Time	Attack Point	MLAD Detect	MLAD Interpret	MLAD Time	MLAD Interpretion
1	28/12/2015 10:29:14	10:44:53	MV-101	-			
2	28/12/2015 10:51:08	10:58:30	P-102	-			
3	28/12/2015 11:22:00	11:28:22	LIT-101	-			
4	28/12/2015 11:47:39	11:54:08	MV-504	-			
5	28/12/2015 11:58:20		No Physical Impact Attack				
6	28/12/2015 12:00:55	12:04:10	AIT-202	+	+	December 28, 2015 12:00:33	AIT202, PITS02, AIT502
7	28/12/2015 12:08:25	12:15:33	LIT-301	+	+	December 28, 2015 12:15:26	LIT301, P101, AIT202
8	28/12/2015 13:10:10	13:26:13	DPIT-301	+	+	December 28, 2015 13:09:52	DPIT301, FIT301, PITS02
9	28/12/2015 14:15:00		No Physical Impact Attack				
10	28/12/2015 14:16:20	14:19:00	FIT-401	+	+	December 28, 2015 14:16:04	FIT401, FIT502, LIT401
11	28/12/2015 14:19:00	14:28:20	FIT-401				
12	29/12/2015 11:10:40		No Physical Impact Attack				
13	29/12/2015 11:11:25	11:15:17	MV-304	+	+	December 29, 2015 11:10:59	MV304, PITS02, MV302
14	29/12/2015 11:35:40	11:42:50	Mv-303	-			
15	29/12/2015 11:52:01		No Physical Impact Attack				
16	29/12/2015 11:57:25	12:02:00	LIT-301	-			
				FP		December 29, 2015 14:33:53	MV101, FIT101, MV303
				FP		December 29, 2015 14:42:44	P602, MV301, MV303
17	29/12/2015 14:38:12	14:50:08	MV-303	-			
18	29/12/2015 18:08:55		No Physical Impact Attack				
19	29/12/2015 18:10:43	18:15:01	AIT-504	-			
20	29/12/2015 18:15:43	18:22:17	AIT-504	+	+	December 29, 2015 18:15:23	AIT504, FIT502, AIT402
21	29/12/2015 18:30:00	18:42:00	MV-101, LIT-101	+	+		
22	29/12/2015 22:55:18	23:03:00	UV-401, AIT-502, P-501	+	+	December 29, 2015 22:54:55	UV401, PITS02, AIT202, AIT501, P205
23	30/12/2015 01:42:34	01:54:10	P-602, DIT-301, MV-302	+	+	December 30, 2015 01:42:11	MV302, DPIT301, PITS02, AIT501
24	30/12/2015 09:51:08	09:56:28	P-203, P-205	+	+	December 30, 2015 09:51:34	P203, P205, P101
25	30/12/2015 10:01:50	10:12:01	LIT-401, P-401	+	+	December 30, 2015 10:11:46	LIT401, MV101, P302
26	30/12/2015 17:04:56	17:29:00	P-101, LIT-301	-			
27	31/12/2015 01:17:08	01:45:18	P-302, LIT-401	+	+	December 31, 2015 01:17:34	LIT401, MV101, AIT501
28	31/12/2015 01:45:19	11:15:27	P-302	-			
				FP		December 31, 2015 11:48:06	MV302, MV304, AIT501
				FP		December 31, 2015 11:48:20	AIT501, P302, AIT402
29	31/12/2015 15:32:00	15:34:00	P-201, P-203, P-205	-			
30	31/12/2015 15:47:40	16:07:10	LIT-101, P-101, MV-201	+	+	December 31, 2015 16:06:31	MV201, LIT101, MV304
				FP		December 31, 2015 22:05:43	MV304, LIT401, P302
31	31/12/2015 22:05:34	22:11:40	LIT-401	+	+	December 31, 2015 22:12:09	LIT401, LIT301, MV301
32	1/01/2016 10:36:00	10:46:00	LIT-301				
33	1/01/2016 14:21:12	14:28:35	LIT-101				
				FP		January 1, 2016 10:37:07	LIT301, MV201, AIT501
				FP		January 1, 2016 10:46:38	LIT301, MV201, LIT401
				FP		January 1, 2016 14:22:21	MV301, LIT101, P602
				FP		January 1, 2016 14:29:25	LIT101, AIT501, AIT402
34	1/01/2016 17:12:40	17:14:20	P-101	+	+	January 1, 2016 17:13:33	MV201, P203, FIT201, P205, P101
35	1/01/2016 17:18:56	17:26:56	P-101; P-102	+	+	January 1, 2016 17:19:59	MV201, P101, FIT101
36	1/01/2016 22:16:01	22:25:00	LIT-101	+	+	January 1, 2016 22:17:27	LIT101, AIT202, AIT402
37	2/01/2015 11:17:02	11:24:50	P-501, FIT-502	+	+	January 2, 2016 11:17:48	FIT504, PITS02, FIT502,
				FP		January 2, 2016 11:28:48	P602, MV303, MV304
38	2/01/2015 11:31:38	11:36:18	AIT-402, AIT-502	+	+	January 2, 2016 11:32:16	AIT502, AIT402, PITS02
39	2/01/2015 11:43:48	11:50:28	FIT-401, AIT-502	+	+	January 2, 2016 11:44:26	FIT401, AIT503, MV101
40	2/01/2015 11:51:42	11:56:38	FIT-401	+	-	January 2, 2016 11:52:19	P602, AIT503, P301
41	2/01/2015 13:13:02	13:40:56	LIT-301	+	+	January 2, 2016 13:41:15	LIT301, LIT401, AIT402

Ex. 3: JSC TANECO - Oil Refining Plant



Facility CDU-VDU-7

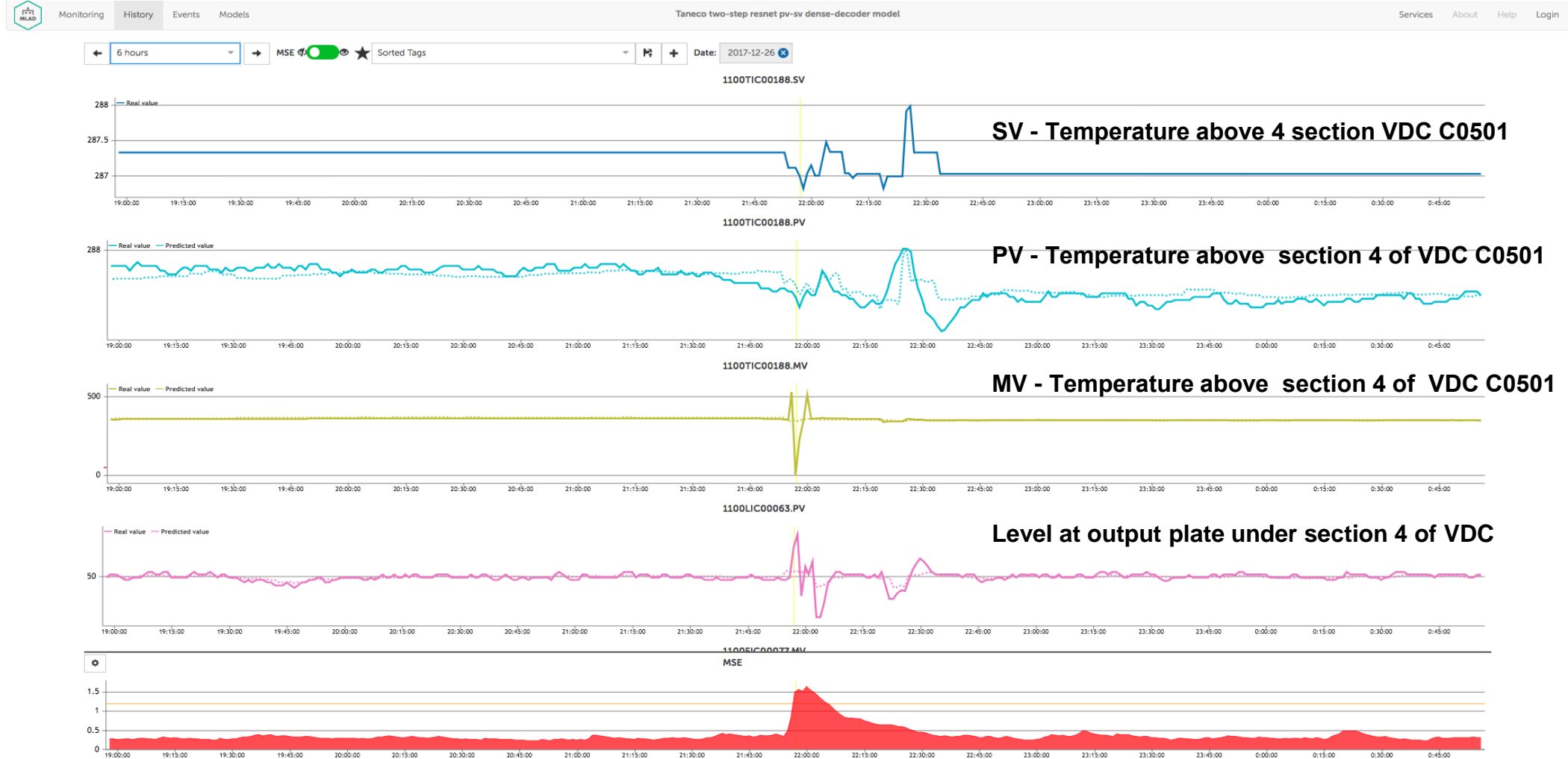
- CDU – crude distillation unit - electric desalting system where crude oil arrives, the furnaces,
- VDU – vacuum distillation unit - the atmospheric and vacuum columns where oil products are distilled into different fractions
- Training on Apr 2017 – Feb 2018 data
- Additional training with new data
- Online monitoring of process anomalies
- Early automatic detection of anomalies in technological processes
- Automatic interpretation of anomalies
- Intuitive interface

Various types of anomalies were detected:

- Deviations in the technological process related to changes in the plant's operating regimes
- Change-over of facility control loops to manual mode
- Situations related to incorrect sensor measurements



Temperature Anomaly in Vacuum Distillation Column (VDC)



Anomaly was caused by an incorrect flowmeter measurement due to impulse tubes being blocked by coke deposits

Distinguishing MLAD Features

MLAD approach

- Noninvasive
- No additional equipment (sensors, non-destructive control equipment, etc.)
- Data-driven anomaly detection
- Detect independently of reason:
 - cyber attack,
 - human factor,
 - equipment faults,....
- Anomaly Interpretation
- Predictive maintenance
- State-of-the-art ML technology
- No need to manually create rules

Other approaches

- Narrow data processing technology capabilities
- Dependent on additional equipment (sensors, non-destructive control equipment, etc.)
- Non-holistic approach
(monitor only certain blocks of plant, doesn't take into account relations with other blocks)

MLAD Value for a Customer

Use Case	Possible Losses	Value
Early Detection <ul style="list-style-type: none">• Quick reaction	<ul style="list-style-type: none">• Non-effectiveness• Outage Losses	↑ of % from process
Automatic Detection <ul style="list-style-type: none">• Known fault situations	Losses from human factor	↑ of % from process ↓ of human factor
Retrospective Analysis <ul style="list-style-type: none">• Processes control effectiveness	Long-lasting non-effective process running	Considerable ↑ of % from process
Anomaly Interpretation <ul style="list-style-type: none">• Processes control effectiveness	Losses from time-consuming manual reason searching	↑ of % from process ↓ of human factor
Seldom Anomaly Detection <ul style="list-style-type: none">• Hard to detect by operator• Serious consequences	Potentially huge: outage, broken equipment, financial loses ∞	Considerable ↓ of risks

Typical MLAD workflow

Stage 0: preliminary study

Prerequisite:

- Industrial technology description
- Acquiring historical data
- Acquiring tag description
- Acquiring technology expertise

Research:

- Data preprocessing
- ML algorithm for data processing
- NN architecture

Delivery:

- ✓ ML-model for historical data
- ✓ Historical Data processing description and results

Stage 1: MLAD piloting

Prerequisite:

- Acquiring on-line data
- Critical tags description
- KICS4Net configuration

Development:

- KICS4Net installation
- Implementation of data processing pipeline
- Implementation and testing of NN model

Delivery:

- ✓ Proof of concept for industrial plant/facility
- ✓ Server with real-time anomaly detection engine

Stage 2: MLAD Implement.

UNDER
DISCUSSION

UNDER
DISCUSSION

UNDER
DISCUSSION

MLAD References

[1] Anomaly Detection for Water Treatment System based on Neural Network with Automatic Architecture Optimization. ICML 2018, DISE1 Workshop, Stockholm, Sweden, 2018

<https://arxiv.org/abs/1807.07282>

[2] RNN-based Early Cyber-Attack Detection for the Tennessee Eastman Process. ICML 2017 Time Series Workshop, Sydney, Australia, 2017.

<https://arxiv.org/abs/1709.02232>

[3] Multivariate Industrial Time Series with Cyber-Attack Simulation: Fault Detection Using an LSTM-based Predictive Data Model. NIPS 2016 Time Series Workshop, Barcelona, Spain, 2016.

<http://arxiv.org/abs/1612.06676>

[4] MLAD: Machine Learning for Anomaly Detection

https://www.youtube.com/watch?time_continue=2&v=1z4nNh9kgbU

<https://ics-cert.kaspersky.com/reports/2018/01/16/mlad-machine-learning-for-anomaly-detection>

[5] MLAD Presentation

https://youtu.be/xXWjfYcPi_Q

mlad@kaspersky.com

Thank you!

mlad@kaspersky.com

