MalwareLab
At the Cyber Security Research Center

Task 10:

Mining Time Oriented Patterns For

Anomaly Detection in ICS

Our Goals and Approach

- Goal:

  Developing an accurate anomaly detection model for ICS based on multivariate time series data (MTSD).

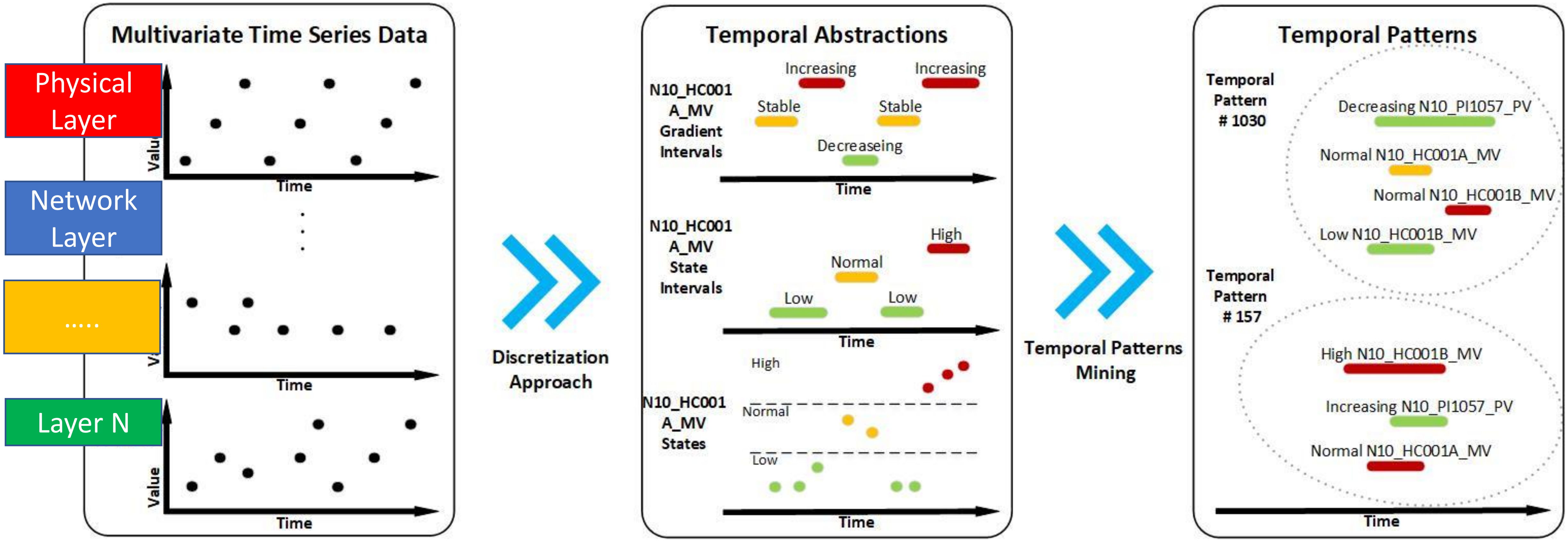
- Our proposed approach:

  - Exploiting and Fusing multiple ICS data sources (Physical layer, Network layer etc.)
  - Mining Time oriented temporal patterns that capture the temporal interaction (B. layers and variables)
  - Induce an ML based detection model that well profile normal ICS behavior over time
  - Detect Anomalous behaviors in ICS based on the profiles we have learned

- Current sub goals:

  - Fully understanding the relevant data that we were provided with (Otorio)
  - Exploring whether the data is enough for our needs
  - Raising our gaps\inputs regarding the data
  - Receiving further data that meets our needs
  - Designing and Developing our proposed detection model based on the updated data we'll receive

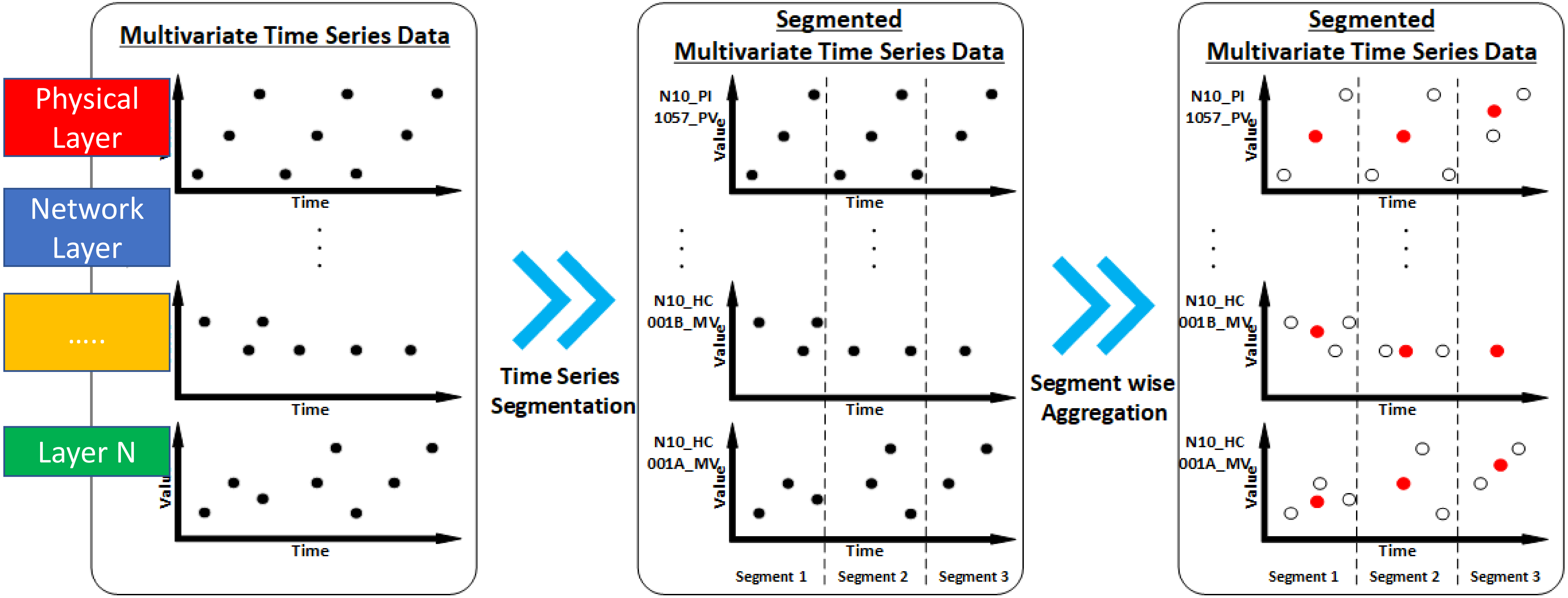
Optimal Case: Temporal Patterns Mining from Multiples Sources



Note: vision can be achieved only if all data sources (layers) will be concurrently recorded from same ICS

Current Data: Temporal Patterns Mining from one source separately

Optimal Case: Temporal Segmentation from Multiples Sources



Note: vision can be achieved only if all data sources (layers) will be concurrently recorded from same ICS

Current Data: Temporal Segmentation from one source separately

MalwareLab At the Cyber Security Research Center

OTORIO - Multivariate Time Series Data (MTSD)

- Full raw logs of IT + OT **(Network + Physical layer)**
- Timestamped data, collected from Meptagon's physical lab environment.
- Including 16 variables derived from the I/O logs of the PLC.
- Recording duration ~5 hours
- Sampling rate ~3Hz
- 37501 Timestamped values derived from the physical layer (the S7-1200 PLC)
- The normal behavior constitutes 86% of the data, while attacks 14%
- Various attacks have been injected into the system

- Our inputs and gaps to be filled:
  - More explanations and descriptions are required regarding the data, especially the attacks conducted including duration and description
  - Data recorded from more **sensors** required to better Profile a Generic Normal Behavior
  - Domain knowledge regarding the values representing a normal behavior
  - Additional data with more attack scenarios
  - More layers of the system (currently data is provided is from one layer, while there are both only)

- We hope OTORIO can assist in addressing those gaps so we can apply our algorithms on it

Task 12:

Temporal Explainability of ICS behavior based on Mined Temporal Segmentations

MalwareLab
At the Cyber Security Research Center

Temporal Segmentation Explainability – Developing Exploration & Visualization Module

- By abstracting the values in each segment, we can achieve an explainability of the behavior across time (segments)
- Using a feature aggregation (features with same functionality) we map a set of features to a specific behavior. For example, different Flow Indicator sensors are related to Flow Indicator

- Flow Indicator – indicating the flow speed in the system pipes. (e.g. in WADI: x , y, z which are the YYY sensors)
- Level Indicator – indicating the water tank(s) level. (e.g. in WADI: x , y, z which are the YYY sensors)
- Pressure Indicator – indicating pressure in the system pipes. (e.g. in WADI: x , y, z which are the YYY sensors)
- Motorized Valve – the level of which the valve is in (out of 3). (e.g. in WADI: x , y, z which are the YYY sensors)

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7	Segment 8	Segment 9	Segment 10
Flow Indicator	Low	Medium	High	High	High	Medium	Medium	Medium	Medium	-
Level Indicator	Low	Medium	Medium	High	-	High	High	Medium	Low	High
Pressure Indicator	Low	Medium	Low	Medium	-	Medium	Medium	Medium	Medium	High

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7	Segment 8	Segment 9	Segment 10
Flow Indicator	-	High	-	Medium	Medium	High	High	Medium	-	Medium
Level Indicator	-	High	-	High	-	High	High	-	High	High
Motorized Valve	High	High	High	Medium	Medium	High	Medium	Medium	-	High

MalwareLab
At the Cyber Security Research Center

Experiments and Current Results

for Tasks 10 and 12

Based on WADI Dataset

WADI - Multivariate Time Series Data (MTSD)

- The WADI dataset (Ahmed et al. 2017) is a water distribution testbed related data.
- WADI consists of a total of five stages:
  - Three stages controlled by Programmable Logic Controllers (PLCs)
  - Two stages controlled via Remote Terminal Units (RTUs).
- The recorded data consists of:
  - 16 days of sampling (14 normal days only ; 2 days containing also attack scenarios)
  - 123 measurements (continuous as well as categorial) regarding the testbed:
    - Actuators (valves etc.) related
    - Sensors (pH etc.) related
  - Sampling rates: 60 Hz (each second)
  - 14 different malicious attacks on different parts (e.g. Sensors, Valves, Pumps) of the testbed
  - High class imbalance – 94% of the data is “no-attack” and 6% is “attack”

WADI - Experimental Design for Classification (Outlier Detection)

MalwareLab
At the Cyber Security Research Center

Main Goal: Comparing between Temporal Patterns mining and Temporal Segmentation Mining in Outlier Detection

Data preprocessing:

- Data splitting according to the shortest attack duration (88 seconds)
- 1,980 samples: 120 are related to 1 of 14 attacks , 1880 samples are "no-attack" (Normal)
  → A class balance of 94% Vs. 6%.

Temporal Segmentation:

- Different number of segments evaluated:
  1, 2, 3, 4, 5 and 10.
- Mean representation of each segment.

Machine Learning Algorithms:

- Variety of ML algorithms; RF, SVM (Linear & RBF kernels), KNN, ANN, NB, and LR.

Goal:

- Evaluate the detection capability of our proposed detection method
- Given new **(unlabeled)** time series of ICS data our method should correctly classify to "attack" or "normal"

Experimental Design:

- The learning methods were evaluated using a stratified 5 folds CV performance was averaged and reported
- Classification performance checked correctness of classifying a given new time series an attack or not.

WADI - Results - Classification (Outlier Detection)

Temporal Patterns Mining

- A total of 121,500 time interval temporal patterns have been discovered
  - A total of 105,000 in the “Attack” class of which around 41,000 are exclusive
  - A total of 80,600 in the “No-Attack” class of which around 16,800 are exclusive
  - A total of 63,700 were mutual for both “Attack” and “No-Attacks”

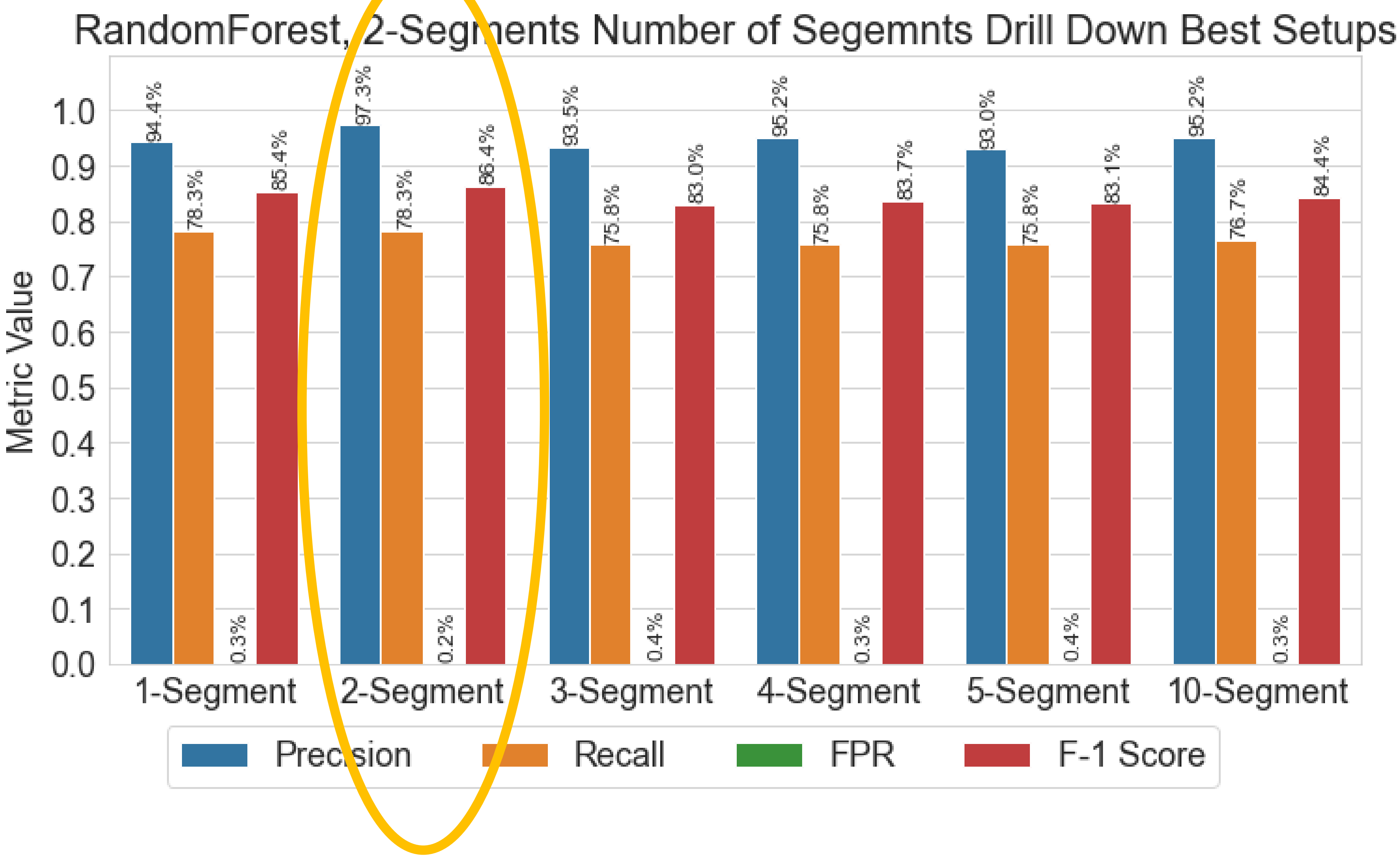
Temporal Segmentation

- 1, 2, 3, 4, 5 and 10 segments evaluated with same time window of 88 seconds

Classification performance (best results are in red) for the task of Outlier Detection

Method	Precision	Recall (TPR)	FPR	F1-Score
KNN (k=5), All TPs, Horizontal Support – Best Precision setup	66.26%	74.43%	18.11%	75.89%
KNN (k=5), Top 100 TPs, Binary, Entropy FS – Best Recall setup	49.21%	87.4%	39.89%	52.1%
KNN (k=5), All TPs, Horizontal Support – Best F1-Score setup	66.26%	74.43%	18.11%	75.89%
RandomForest, 2-Segments – Best setup	97.31%	78.33%	0.16%	86.28%

WADI - Results - Classification (Outlier Detection)



MalwareLab At the Cyber Security Research Center

WADI - Experimental Design for Anomaly Detection (Novelty Detection)

Main Goal: Comparing between Temporal Segmentation and SOTA (MADGAN) in Novelty Detection

Data preprocessing:

- Data splitting according to the shortest attack duration (88 seconds)
- 1,980 samples: 120 are related to 1 of 14 attacks , 1880 samples are "no-attack" (Normal)
  → A class balance of 94% Vs. 6%.

Temporal Abstraction & Temporal Patterns mining:

- State abstraction (only) using Equal Frequency Discretization (EFD)
- A vertical support of 50%
- Mining Temporal Patterns of up to size 3 (including)

Machine Learning Algorithms:

- Feature representation using Horizontal Support , Binary
- Feature selection using: Entropy, Gini;
- Selecting different amounts of temporal patterns: 25, 50, 100, 200, 300, 400, 500 and All
- Variety of ML algorithms; RF, SVM (Linear & RBF kernels), KNN, ANN, NB, and LR.

Goal:

- Evaluate the detection capability of our proposed detection method
- Given new **(unlabeled)** time series of ICS data our method should correctly classify to "attack" or "normal"

Experimental Design:

- The learning methods were evaluated using a stratified 5 folds CV performance was averaged and reported
- Classification performance checked correctness of classifying a given new time series an attack or not.

WADI - Results – Anomaly Detection (Novelty Detection)

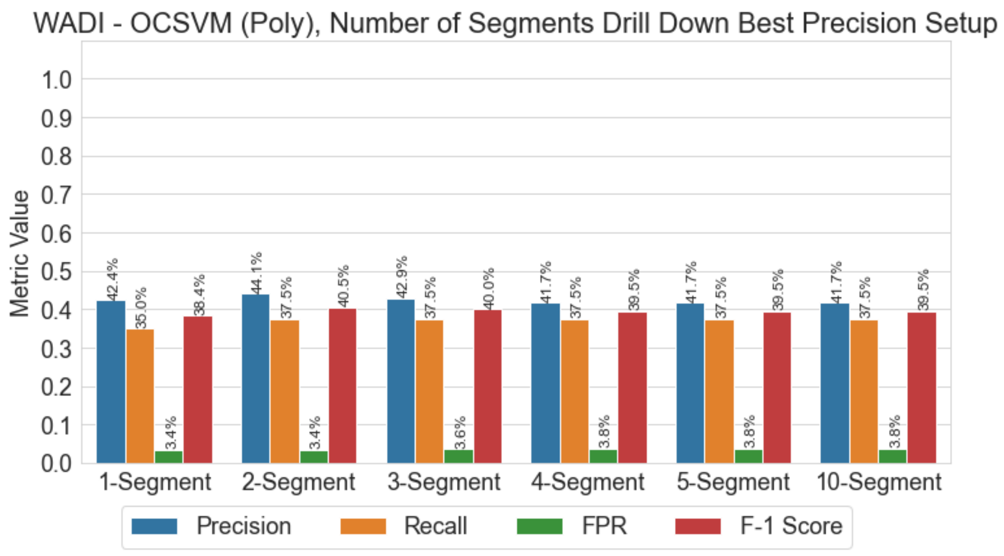
Temporal Segmentation

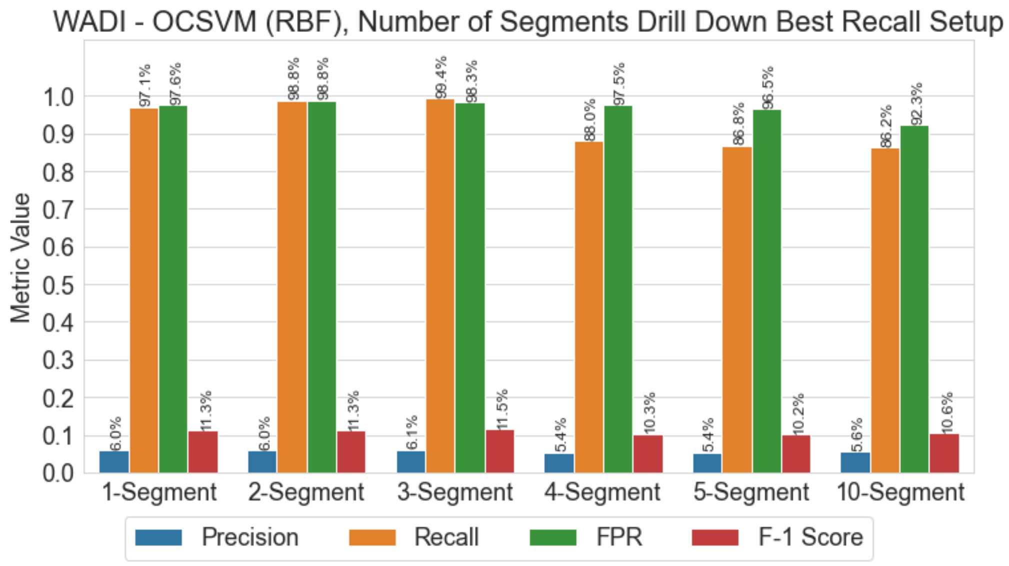
- 1, 2, 3, 4, 5 and 10 segments evaluated
- Sliding window of 30, 60, 90, 120, 150, 180, 210, 240, 270 and 300 seconds

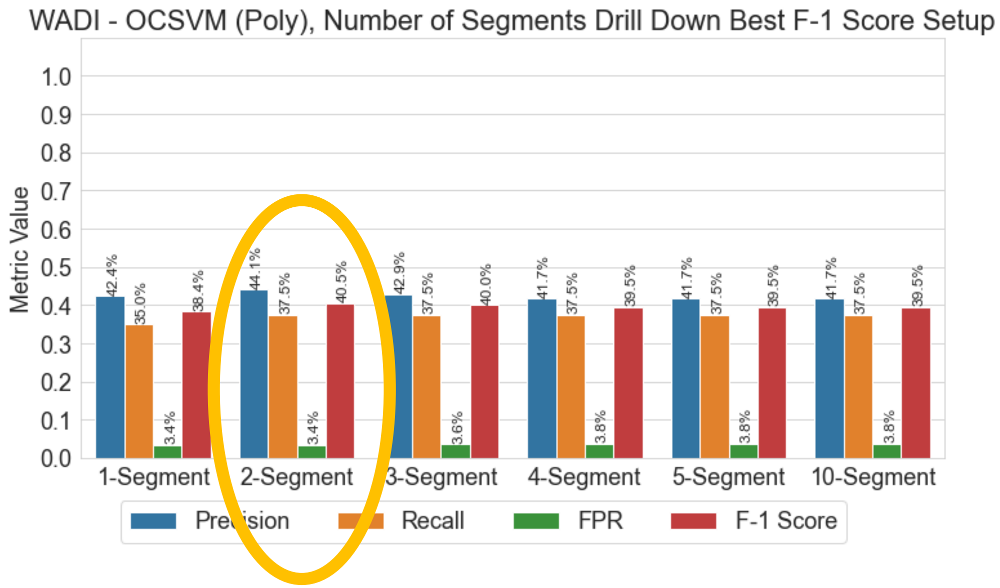
Anomaly detection performance (best results are in red)

Method	Precision	Recall (TPR)	FPR	F1-Score
OCSVM (Poly), 2-Segments, 300s Window – Best Precision setup	44.11%	37.5%	3.4%	40.5%
OCSVM (RBF), 3-Segments, 60s Window – Best Recall setup	6.1%	99.99%	98.3%	11.5%
OCSVM (Poly), 2-Segments, 300s Window – Best F1-Score setup	44.11%	37.5%	3.4%	40.5%
MAD-GAN (Li et al. 2019) – Best Precision setup	46.98%	24.58%	NA	32%
MAD-GAN (Li et al. 2019) – Best Recall setup	6.46%	99.99%	NA	12%
MAD-GAN (Li et al. 2019) – Best F1-Score setup	41.44%	33.92%	NA	37%

WADI - Segment Size Drill Down Results







MalwareLab
At the Cyber Security Research Center

WADI - Temporal Segmentation Explainability

Highlighting and visualizing segments time, one can easily explain temporal behaviors.

Attack

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7	Segment 8	Segment 9	Segment 10
Flow Indicator	Low	Low	-	Medium	Medium	High	Medium	Low	Low	High
Level Indicator	-	Low	Low	Low	Low	Low	Low	Low	Low	High
Pressure Indicator	Medium	Medium	Medium	Medium	Medium	Medium	Low	Low	-	High

No-Attack

	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7	Segment 8	Segment 9	Segment 10
Flow Indicator	Low	Medium	High	High	High	Medium	Medium	Medium	Medium	-
Level Indicator	Low	Medium	Medium	High	-	High	High	Medium	Low	High
Pressure Indicator	Low	Medium	Low	Medium	-	Medium	Medium	Medium	Medium	High

Cooperation and Commercialization

- Initial Cooperation With OTORIO – based on the data they have provided

- Commercialization – once we have more data, we will be able to better understand the relative advantage of our proposed solutions and its commercialization possibilities.