

Task 10:

Mining Time Interval Temporal Patterns For Anomaly Detection in ICS

PI: Dr. Nir Nissim (BGU)

Our Goals and Approach

- Goal:

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

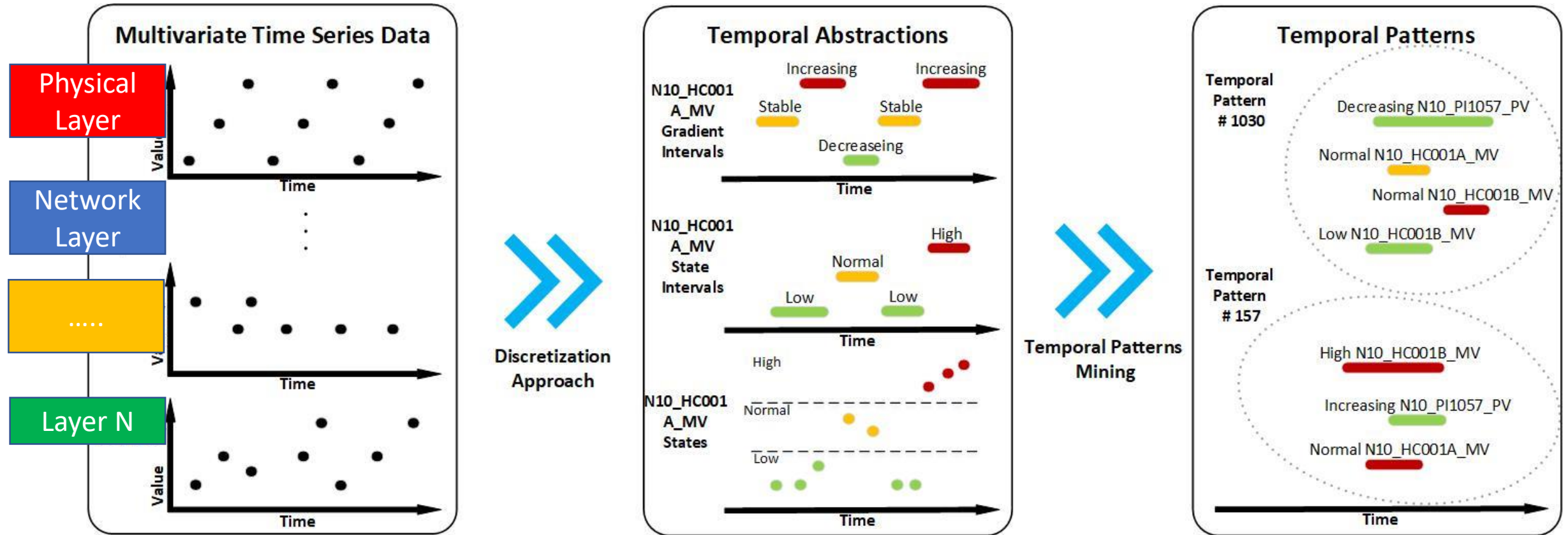
- Our proposed approach:

- Exploiting and Fusing multiple ICS data sources (Physical layer, Network layer etc.)
- Mining Time Interval temporal patterns
- Induce an ML based detection model that well profile normal ICS behavior over time
- Detect Anomalous behaviors in ICS based on

- Current sub goals:

- Fully understanding the data that we were provided with (Delek, Otorio, DLC)
- 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 concurrent recorded of the same ICS

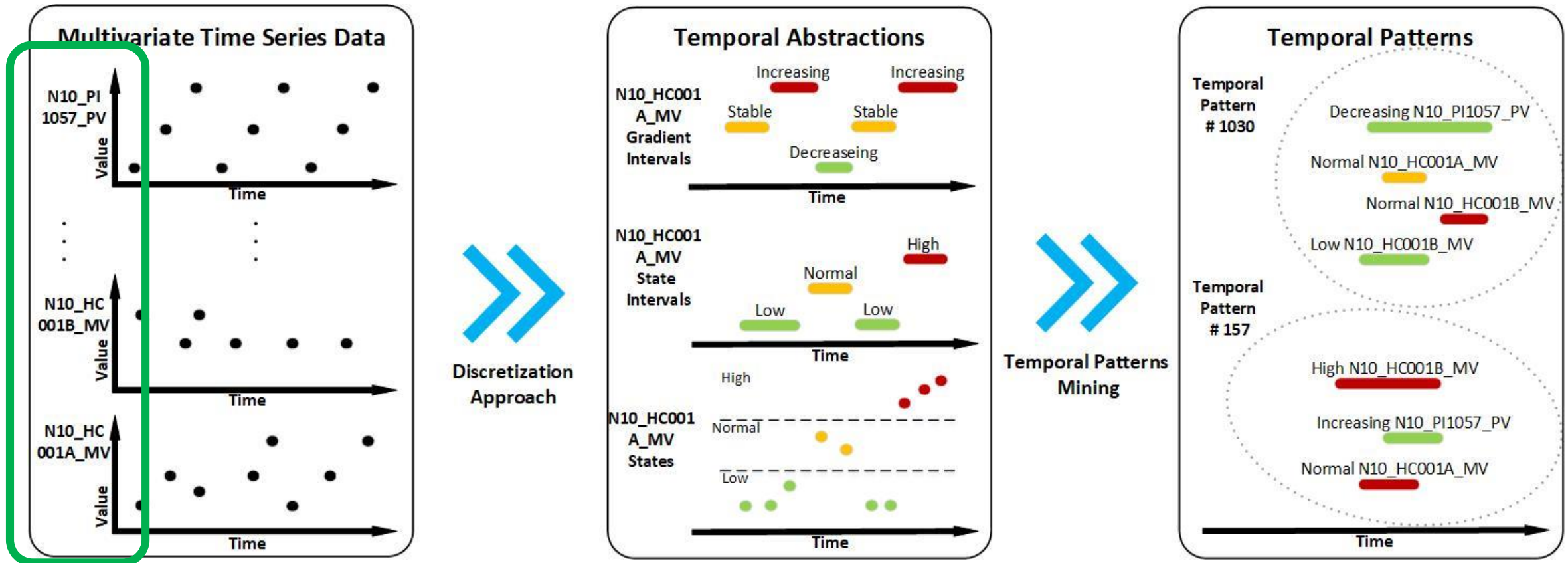
Current Data: Temporal Patterns Mining from one source each

Delek-US PI - Multivariate Time Series Data (MTSD)

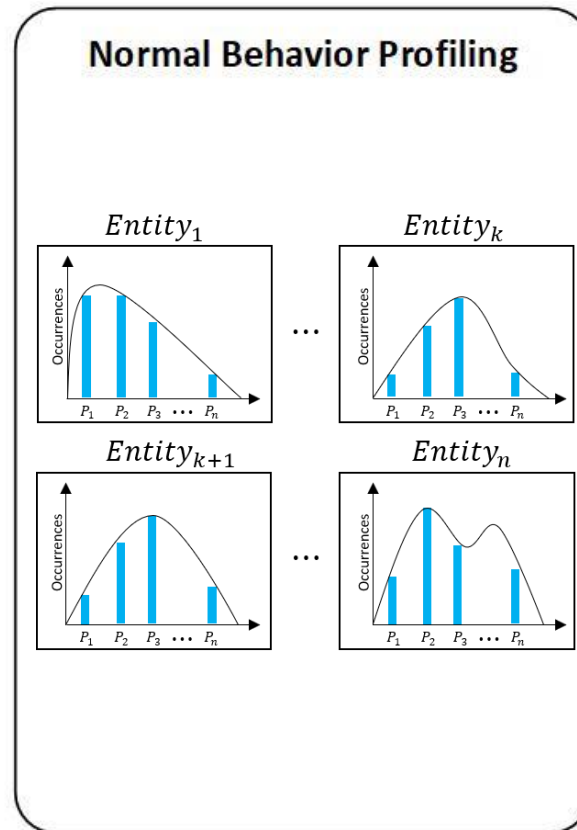
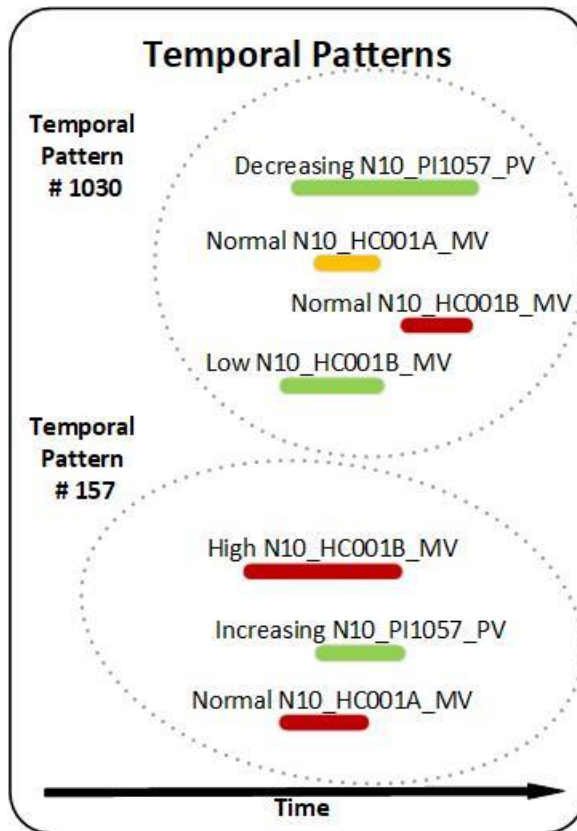
- A Recording of a Single ICS entity (Physical layer)
- Timestamped data, collected from hundreds of sensors via the PI framework 556 Raw Features (different sensor)
- Recording duration ~12 months
- Sampling rate every 30 minutes
- 17521 Timestamped values (in total from all data sources\ sensors)
- Our inputs and gaps to be filled:
 - More explanations and descriptions are required regarding the data (Bob?)
 - Higher Sampling rate is required (much lesser than 30 current minutes, e.g. every minute)
 - Data recorded from more ICS entities is required to better Profile a Generic Normal Behavior
 - Malicious or Anomalous data should be provided to evaluate the model

	DateTime	DateTime_Elapsed	DateTime_year	DateTime_month	DateTime_day	DateTime_hour	DateTime_minute	DateTime_second	DateTime_weekday	N10_HC001A_MV	...	N10_FI1053_PV
0	2860.39990	0.000000	2017	10	31	9	35	45	3	58.999939	...	5.422828
1	2860.42065	0.020833	2017	10	31	10	5	45	3	58.999939	...	5.188814
2	2860.44141	0.041667	2017	10	31	10	35	45	3	58.999939	...	4.270872
3	2860.46240	0.062500	2017	10	31	11	5	45	3	58.999939	...	5.009398
4	2860.48315	0.083333	2017	10	31	11	35	45	3	58.999939	...	4.636016

Temporal Patterns Mining – Delek-US MTSD Example



Temporal Patterns For Anomaly Detection – Delek-US MTSD Example



Anomaly Detection

	<i>Test Entity₁</i>	<i>Test Entity₂</i>	<i>Test Entity₃</i>
<i>Entity₁</i>	0.5	0.2	0
<i>Entity₁</i>	0.1	0.3	0.9
<i>Entity₁</i>	0.95	0.8	0.2
<i>Entity₁</i>	0.1	0.6	0.5
<i>Entity₁</i>	0.77	0.55	0
<i>Entity₁</i>	0.34	0.05	0.23

OTORIO - Multivariate Time Series Data (MTSD)

- Recording of network data of 10 machines (Network layer)
- Timestamped data of 3 raw features (captured packed length; packet length; and the packet data)
- Recording duration 263 seconds
- Sampling rate - varied
- 969 timestamped values (in total from all data sources\sensor)
- Our inputs:
 - More features are required (3 are not enough)
 - Malicious or Anomalous data should be provided to evaluate the model

	interface_id	timestamp_high	timestamp_low	captured_len	packet_len	packet_data
0	0.0	384093	2927823909	98	98	b"\x00\t\x0f\t\x00\x04\xe8\x84\xa5\x18\x95\x07...
1	0.0	384093	2927823909	98	98	b"\x00\t\x0f\t\x00\x04\xe8\x84\xa5\x18\x95\x07...
2	0.0	384093	2927823932	98	98	b"\x00\t\x0f\t\x00\x04\xe8\x84\xa5\x18\x95\x07...
3	0.0	384093	2927824003	98	98	b"\x00\t\x0f\t\x00\x04\xe8\x84\xa5\x18\x95\x07...
4	0.0	384093	2927824082	98	98	b"\x00\t\x0f\t\x00\x04\xe8\x84\xa5\x18\x95\x07...

DLC - Univariate Time Series Data (a non MTSD)

- Recording of Smart Meter Data (probably physical layer)
- Univariate timestamped data of electrical consumption
- Recording duration **1** month (for both 15kwh and 60 kwh data collections)
- Sampling rate every 15 minutes (for 15kwh data) and 1 hour (for 60kwh data)
- 146,899,065 Timestamped values (for 15kwh data)
- 40,386,347 Timestamped values (for 60kwh data)
- Our inputs:
 - More features are required (one is not enough)
 - Malicious or Anomalous data should be provided to evaluate the model

	MTR_CONFIG_TY_CD	MTR_INSTALL_DTTM	MTR_REMOVAL_DTTM	CITY	STATE	ZIP	MEASR_COMP_USAGE_FLG	MSRMT_LOCAL_DTTM	MSRMT_VAL
0	AMI1P-15	2017-02-21 15:03:50	2019-07-03 07:59:25	PITTSBURGH	PA	15209	+	2019-06-01 01:00:00	0.169000000000000000
1	AMI1P-15	2017-02-21 15:03:50	2019-07-03 07:59:25	PITTSBURGH	PA	15209	+	2019-06-01 01:15:00	0.242000000000000000
2	AMI1P-15	2017-02-21 15:03:50	2019-07-03 07:59:25	PITTSBURGH	PA	15209	+	2019-06-01 01:30:00	0.148000000000000000
3	AMI1P-15	2017-02-21 15:03:50	2019-07-03 07:59:25	PITTSBURGH	PA	15209	+	2019-06-01 01:45:00	0.143000000000000000
4	AMI1P-15	2017-02-21 15:03:50	2019-07-03 07:59:25	PITTSBURGH	PA	15209	+	2019-06-01 02:00:00	0.157000000000000000