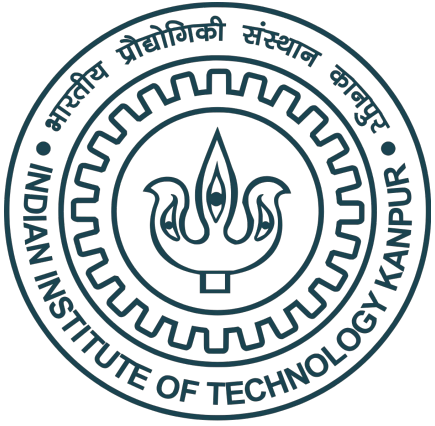
Anomaly Detection for Cyber-Physical System

(CS631A Project)



Presented To:

Dr. Sandeep Shukla
Professor, CSE Dept.
IIT Kanpur, UP

Presented By:

Team - Bob the Builder
Members - Nanda Rani (21111265)
Trishie Sharma (21111270)
Harshvardhan Pratap Singh (20111410)
Harsika Diksha (20111021)
Nikhil Molugu (20111415)

Index

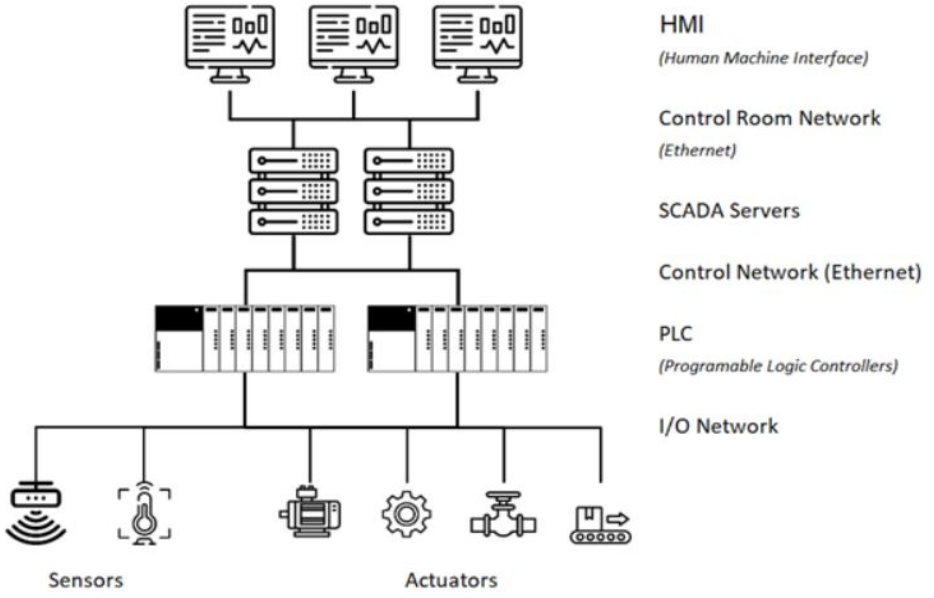
- Introduction
- Threat Model
- Objective
- Dataset
- Methodology
- Result & Discussion
- Conclusions

Introduction

- Cyber-Physical System (CPS) is collection of networking, computation and physical processes
- Unlike traditional IT systems, which mainly manage data, ICS control physical processes.
- Impact of cyber attacks is no longer bounded by financial losses.
- Growing interest in developing techniques that can detect such sophisticated attacks at the process level.
- A process-level intrusion detection system monitors sensors and possibly control commands, to determine if the physical process is drifting from the normal or expected behavior.
- Ofcourse, we cannot forget Stuxnet [4].

Threat Model

- The communication link between the PLC and SCADA.
- APT Malware threat, Insider threat and Dos attack.
- Network based attack can also change dynamics of the system.
- Need to have some continuous monitoring system to monitor continuous system behaviour.



Overview of ICS system [6]

Objective

- To develop detection mechanism for anomalies detection for Cyber-Physical System using TE-dataset.
- Implementation of some state of the art solutions for anomalies detection such as PASAD and M-PASAD.
- Comparative analysis of implemented models.

Dataset

- Based on Tennessee-Eastman process control model [5].
- TE-dataset [1] has approx 4800 measurements of 41 sensors.
- Prepared by simulating integrity attacks on both actuators and sensors.
  - Stealth Attack
  - Direct damage Attack
- 100 measurement per hour and 40 hr normal behaviour and 8 hr under attack measurements.

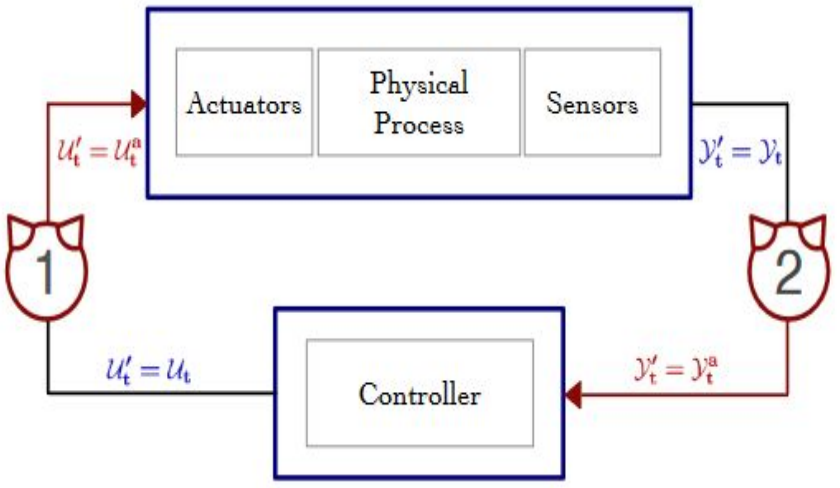
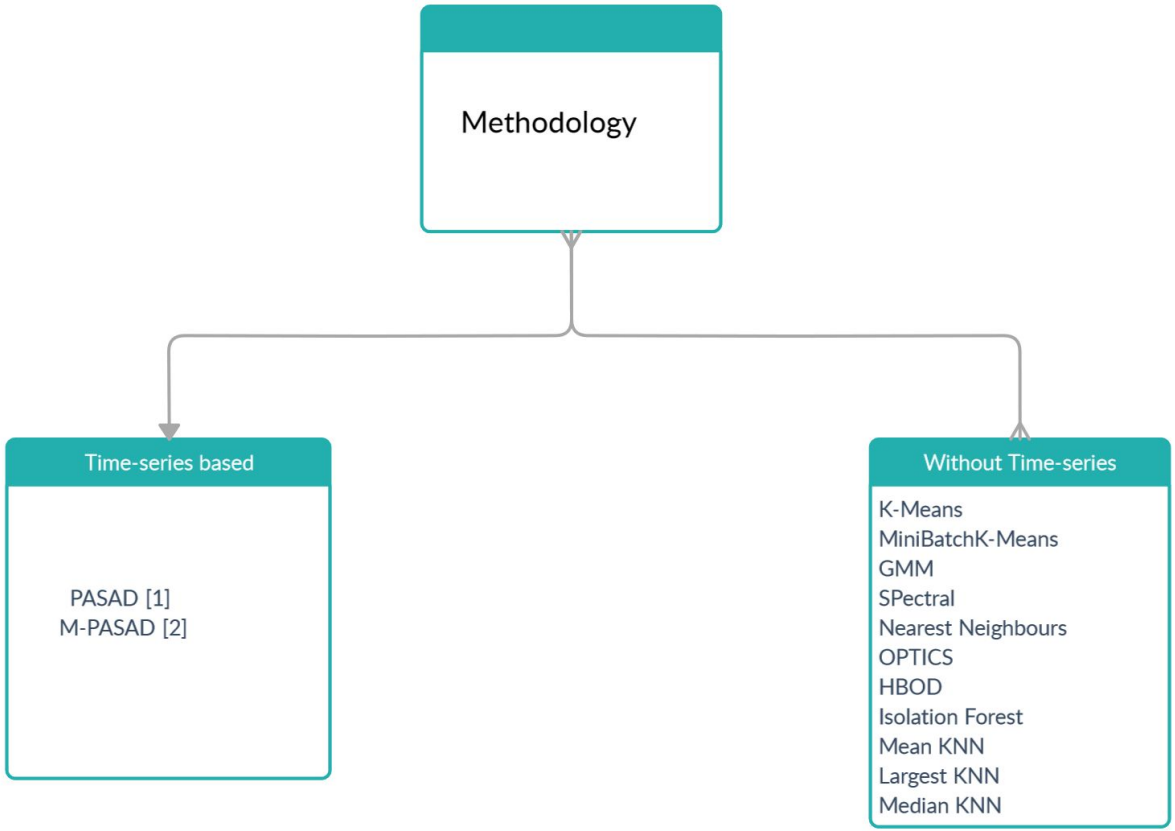


Fig: Attack scenarios on control systems: Attacks on actuator signals (1) and attacks on sensor signals (2). [1]

Methodology



PASAD

- Process-Aware Stealthy Attack Detection [1]
- Trained on normal behaviour and check whether new observations are departing the normal behaviour or not.
- Univariate detection system
- Proved Isometric trick

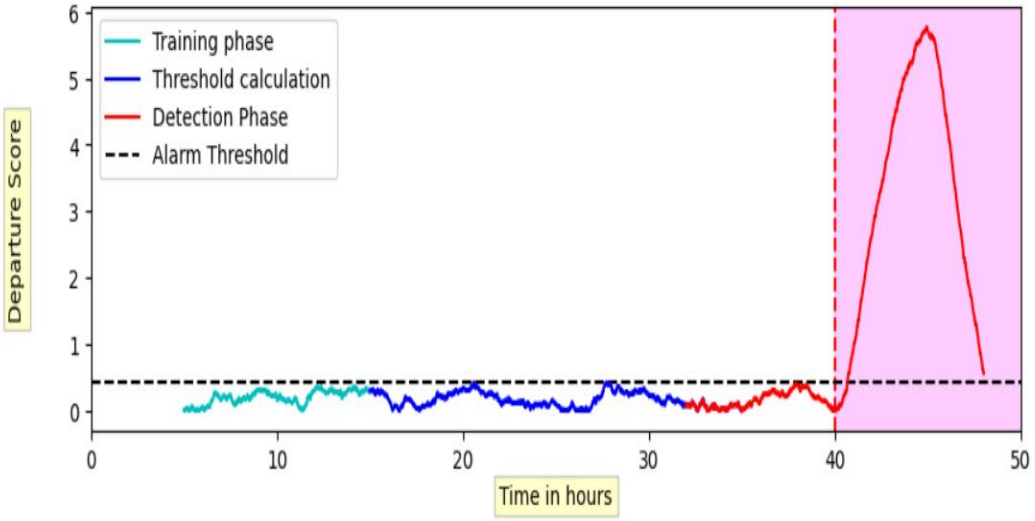
$$||\mathbf{U}^T\mathbf{x}|| = ||\mathbf{U}\mathbf{U}^T\mathbf{x}||$$

- Distance calculated by euclidean distance between the most recent test vector and the centroid of the cluster.
- Centroid: $\tilde{\mathbf{c}} = \mathbf{P}\mathbf{c}$

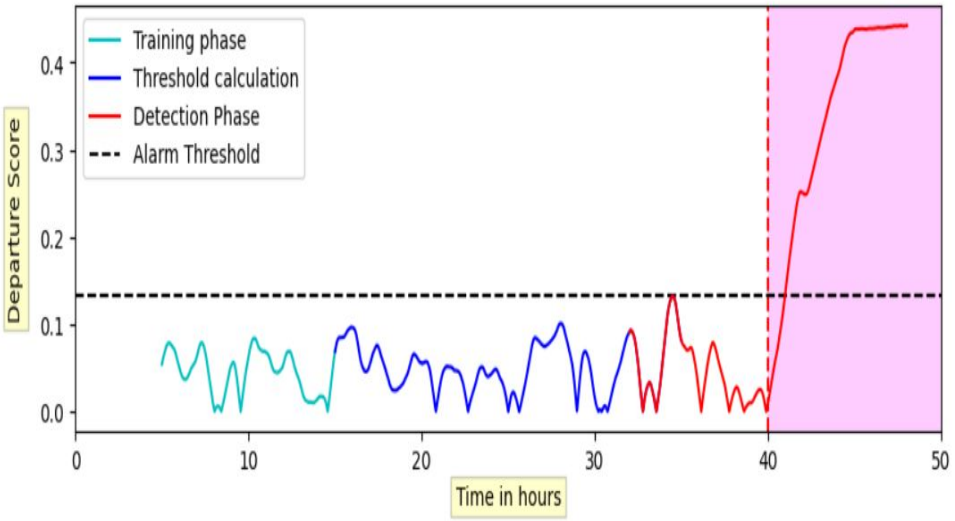
  Where c is mean of sample and P is projection Matrix.

Embeddings	• Trajectory matrix from lag vector • Size: L * N
Singular Value Decomposition	• Obtain eigenvector and eigenvalues of lag covariance matrix • Calculate value for r
Projection onto Signal Subspace	• Evaluate U (L*r) • Calculate centroid.
Distance Tracking	• $Dj = ||\hat{c} - Pxj||$ • $Dj \geq \boldsymbol{\theta}$

PASAD Result



PASAD can detect Direct damage attacks



PASAD can detect Stealth attacks

M-PASAD

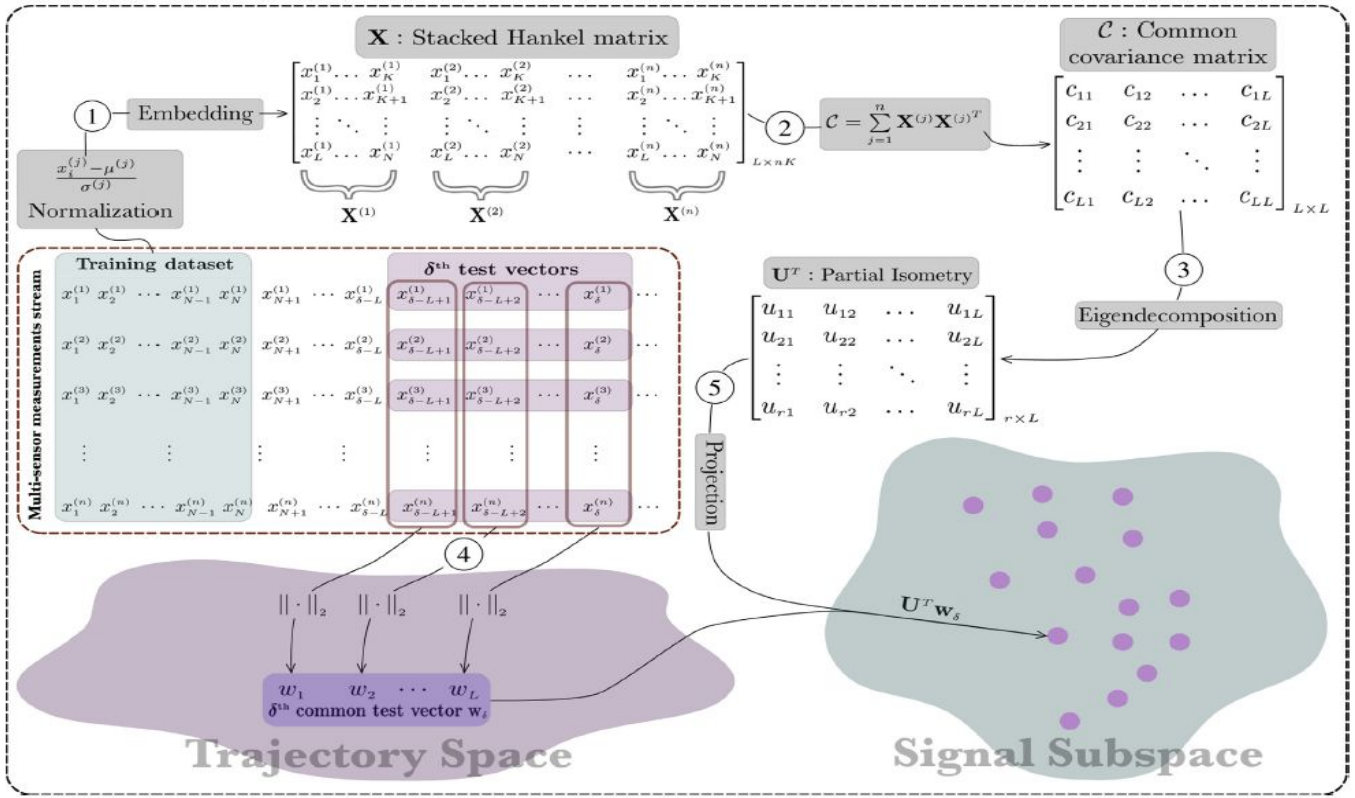
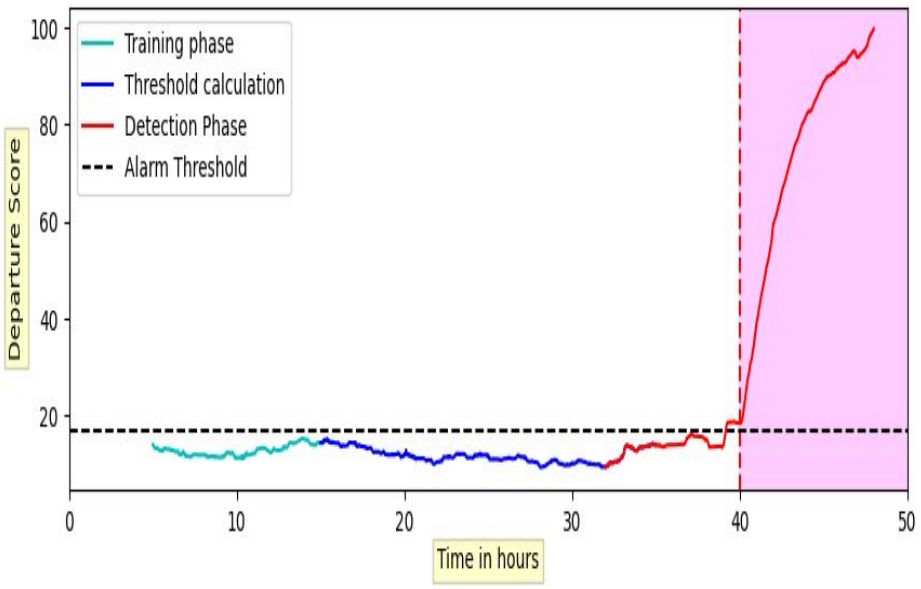
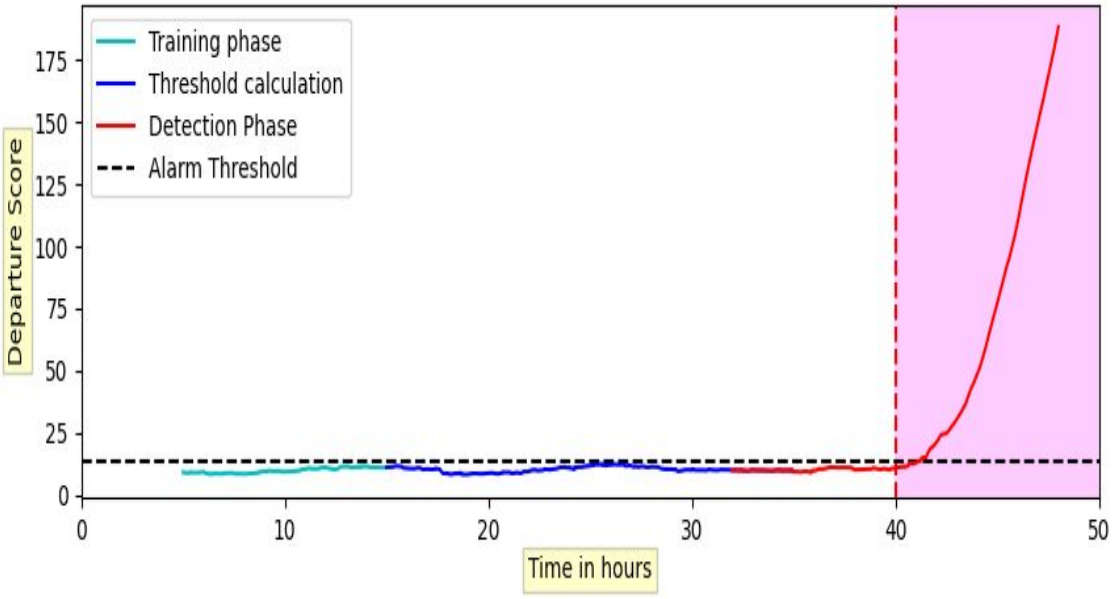


Fig: Workflow of M-PASAD

M-PASAD Result



M-PASAD can detect Stealth attacks



M-PASAD can detect Direct damage attacks

K Means Clustering

- Search for a predetermined number of clusters within an unlabeled multidimensional dataset.
- It accomplishes this using a simple conception of what the optimal clustering looks like:
  - The "cluster center" is the arithmetic mean of all the points belonging to the cluster.
  - Each point is closer to its own cluster center than to other cluster centers.

Algorithm:

- Find clusters using K-Means algorithm with normal data(starting 3200 rows).
- Calculate Threshold
  a. Take distance between point and centroid with maximum distance by avoiding few percent outliers.
- Calculate distance between test point and centroid of clusters.
  a. If(distance >= Threshold):
  b. Point is anomaly
- Calculate Accuracy with 50:50 ratio of normal and under attack data points.

Gaussian Mixture Model

- Attempt to find a mixture of multi-dimensional Gaussian probability distributions that best model any input dataset.
- Can be used for finding clusters in the same manner as *k-means* But because GMM contains a probabilistic model under the hood, it is also possible to find probabilistic cluster assignments.

Algorithm:

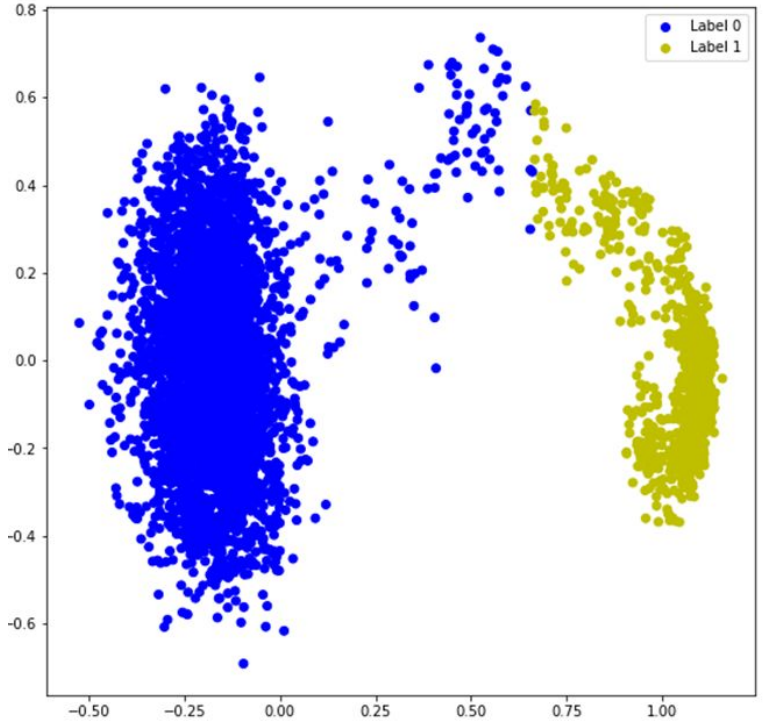
- Find clusters with two components using GMM algorithm complete data(4800 data pointsof each 4 data files). It learned under hood probability distribution of both Natural and under attack situation.
- Tested the algorithm with under attack data points of other files by finding probabilities.
- If Probability is more than 0.5 in whichever cluster it will belong to that cluster.
- Accuracy is calculated with other files under attack situations.

Spectral Clustering

- Relies on the power of graphs and the proximity between the data points to cluster them, unlike K-means where sphere shape cluster are always assumed.
- PCA-2 dimensionality reduction.

Algorithm:

- First, we construct a nearest neighbors graph (KNN graph) or radius based graph.
- Then embed the data points in low dimensional space (spectral embedding) in which the clusters are more obvious with the use of eigenvectors of the graph Laplacian.
- Then, use the lowest eigenvalue to choose the eigenvector for the cluster.

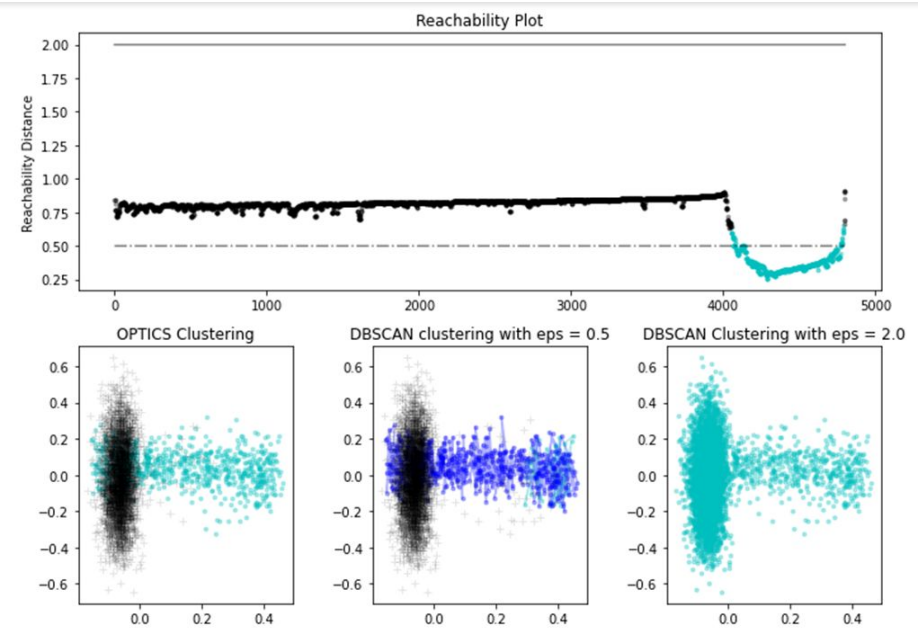


OPTICS Clustering

- This algorithm works on creating a reachability plot which is used to extract clusters.

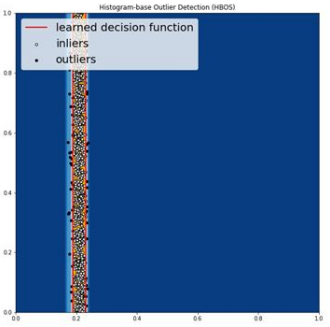
Algorithm:

- First, we start out by calculating the core distances on all data points in the set.
- Then we will loop through the entire data set, and update the reachability distances, processing each point only once, trying to keep clusters near to each other.
- Next, we extract the actual cluster labels from the plot by searching for “valleys” in the plot, using local minimums and maximum.
- The cluster score is calculated on the predicted labels of the clusters.

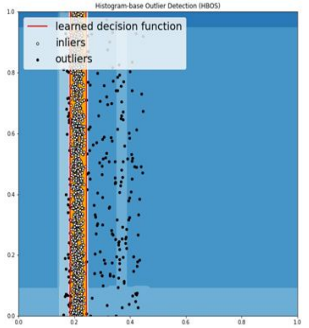


Histogram-based Outlier Detection

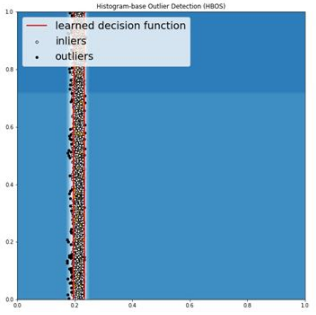
- It is a fast unsupervised method that assumes feature independence and calculates outlier scores by constructing histograms.
- The algorithm presupposes that the characteristics are totally independent of one another (zero multicollinearity).
- Univariate – Frequencies or count of the distinct classes
- Multivariate – Sum of individual histogram density of each feature



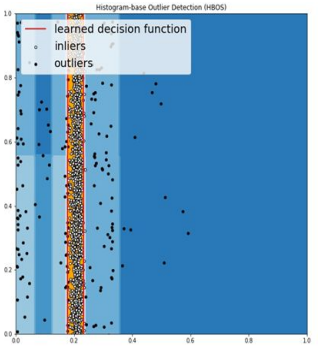
SA1



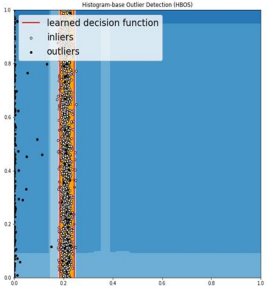
SA3



SA2



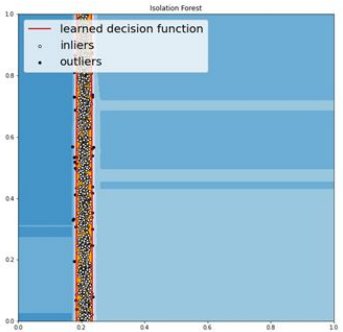
DA1



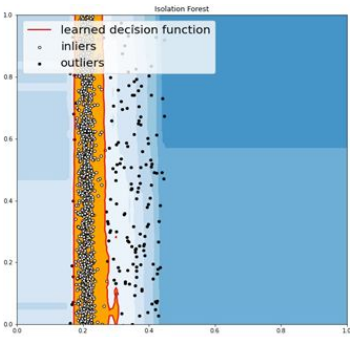
DA2

Isolation Forest

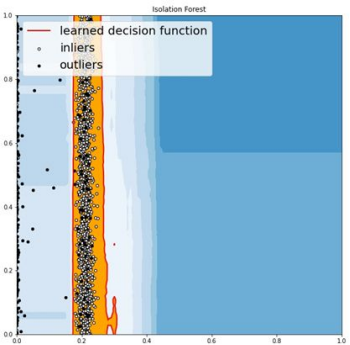
- It employs the scikit-learn library and performs data partitioning via a set of decision trees.
- It calculates an anomaly score based on how isolated a certain location is within the structure.
- The anomaly score is then used to distinguish outliers from the rest of the data.
- It is observed that Isolation Forest performs well with multi-dimensional data.



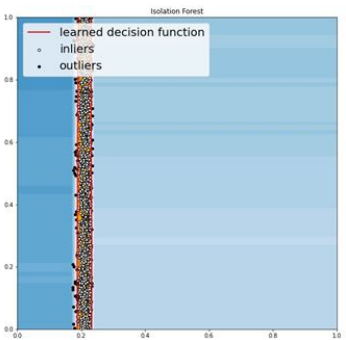
SA1



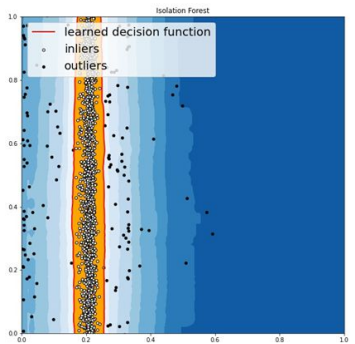
SA3



DA2



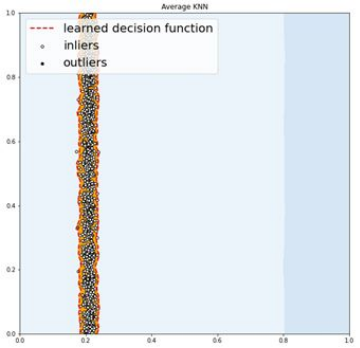
SA2



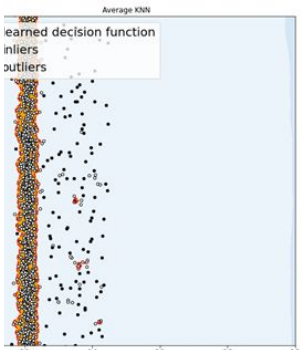
DA1

k-Nearest Neighbors

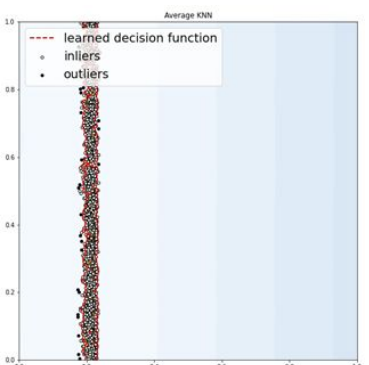
- The outlying score for any data point could be defined as the distance to its k^{th} nearest neighbor and kNN algorithm gives the output using this outlying score.
- There are three variants of kNN used in this work:
  - Largest kNN : The outlier score is calculated using the distance of the k^{th} neighbour.
  - Mean kNN : The outlier score is calculated using the average of all k neighbours.
  - Median kNN : The outlier score is calculated using the median of the distance to k neighbours.



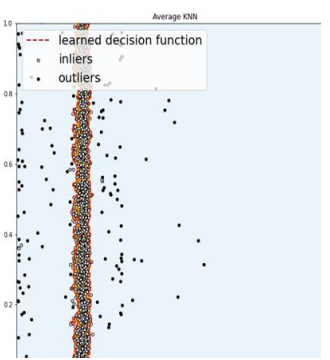
SA1



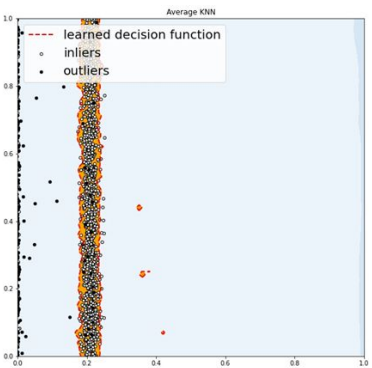
SA3



SA2

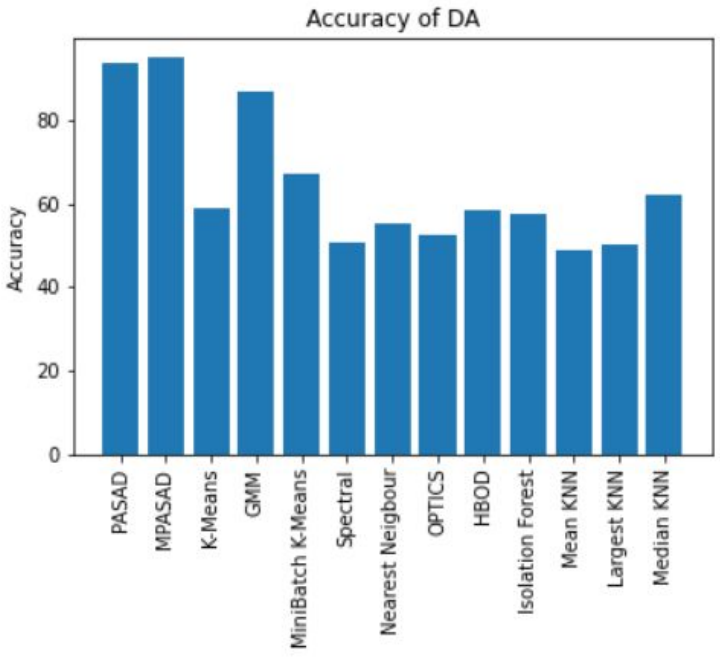


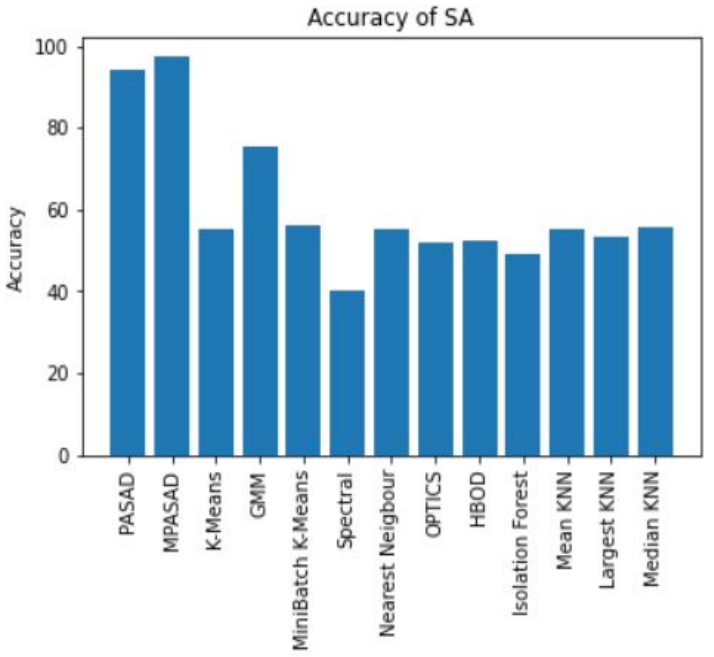
DA1



DA2

Results & Discussion





- M-PASAD has performed well for both types of attack.

Result & Discussion

M-PASAD Performance Metric

	Precision	Recall	f1-score	Accuracy
SA1	0.89	1.0	0.94	94.88
SA2	1.0	0.73	0.84	82.27
SA3	1.0	0.94	0.97	97.31
DA1	1.0	0.65	0.79	73.59
DA2	1.0	0.86	0.92	92.34

Time taken for detection by M-PASAD = 12.84 ms
Time-taken for detection by PASAD = 24.40 ms

Time complexity of PASAD is approx 1.9 times time taken by M-PASAD

Conclusions

- M-PASAD (time-series based) method found to be best among all implemented methodology.
- Capable of detecting sophisticated attacks by monitoring time series of sensor measurements for structural changes in their behavior.
- Less computational cost such as less space and time complexity
- Unlike PASAD, M-PASAD is a multivariate.
- Time-series aspects of data cannot be ignored.
- M-PASAD can detect both Stealthy and Direct damage attack and having best accuracy tends to 98% on our chosen dataset.

Reference

[1] W. Aoudi, M. Iturbe, M. Almgren, Truth will out: departure-based process-level detection of stealthy attacks on control systems, in: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS '18, ACM, New York, NY, USA, 2018, doi: 10.1145/3243734.3243781 .

[2] Wissam Aoudi and Magnus Almgren. A scalable specification-agnostic multi-sensor anomalydetection system for iiot environments.Int. J. Crit. Infrastructure Prot., 30:100377, 2020.

[3] H. Hassani, R. Mahmoudvand, Multivariate Singular Spectrum Analysis, Pal- grave Macmillan UK, London, pp. 49–86. 10.1057/978-1-137-40951-5_2

[4] Thomas Chen and Saeed Abu-Nimeh. 2011. Lessons from Stuxnet. Computer (2011).

[5] James Downs and Ernest Vogel. A plant-wide industrial process control problem.Computers& Chemical Engineering, 1993.

[6] Patrick Berry. Processing SCADA Alarm Data Offline with ELK. (2021) online Available at: https://towardsdatascience.com/processing-scada-alarm-data-offline-with-elk-7ab9b475ffb9

Thank You