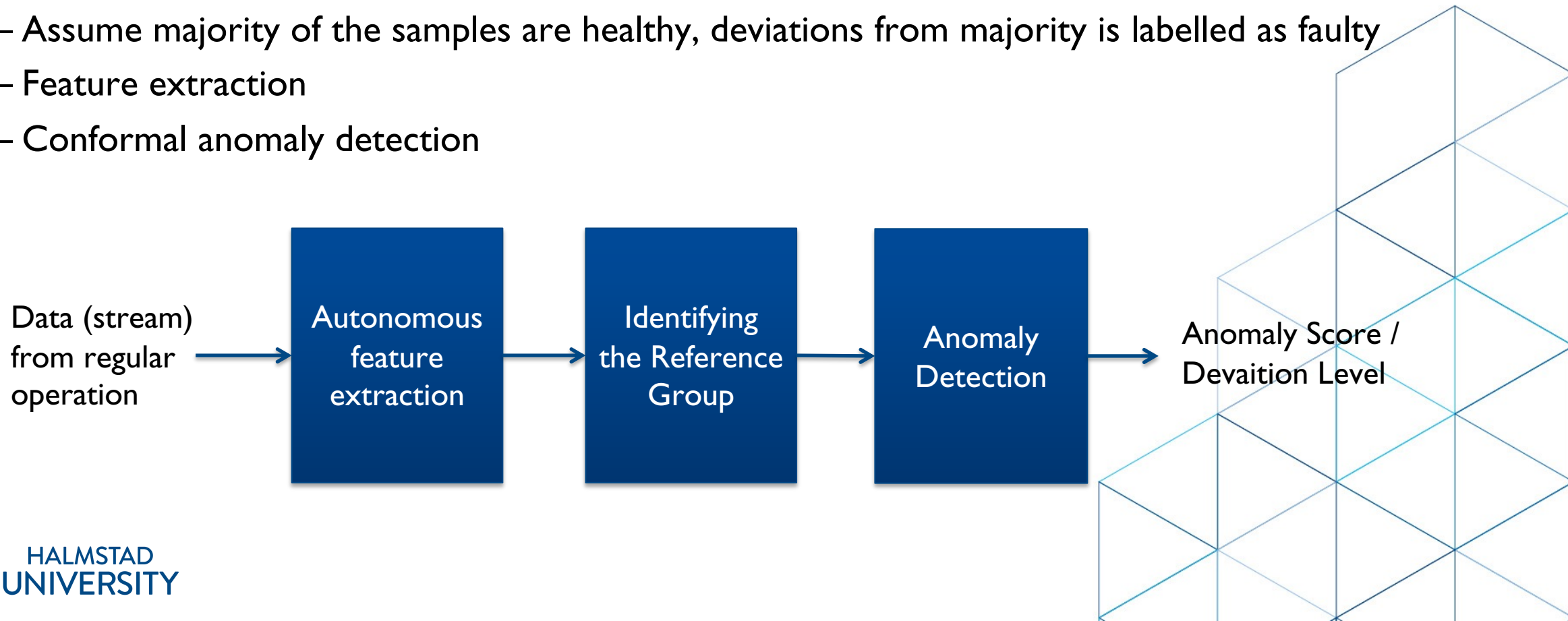


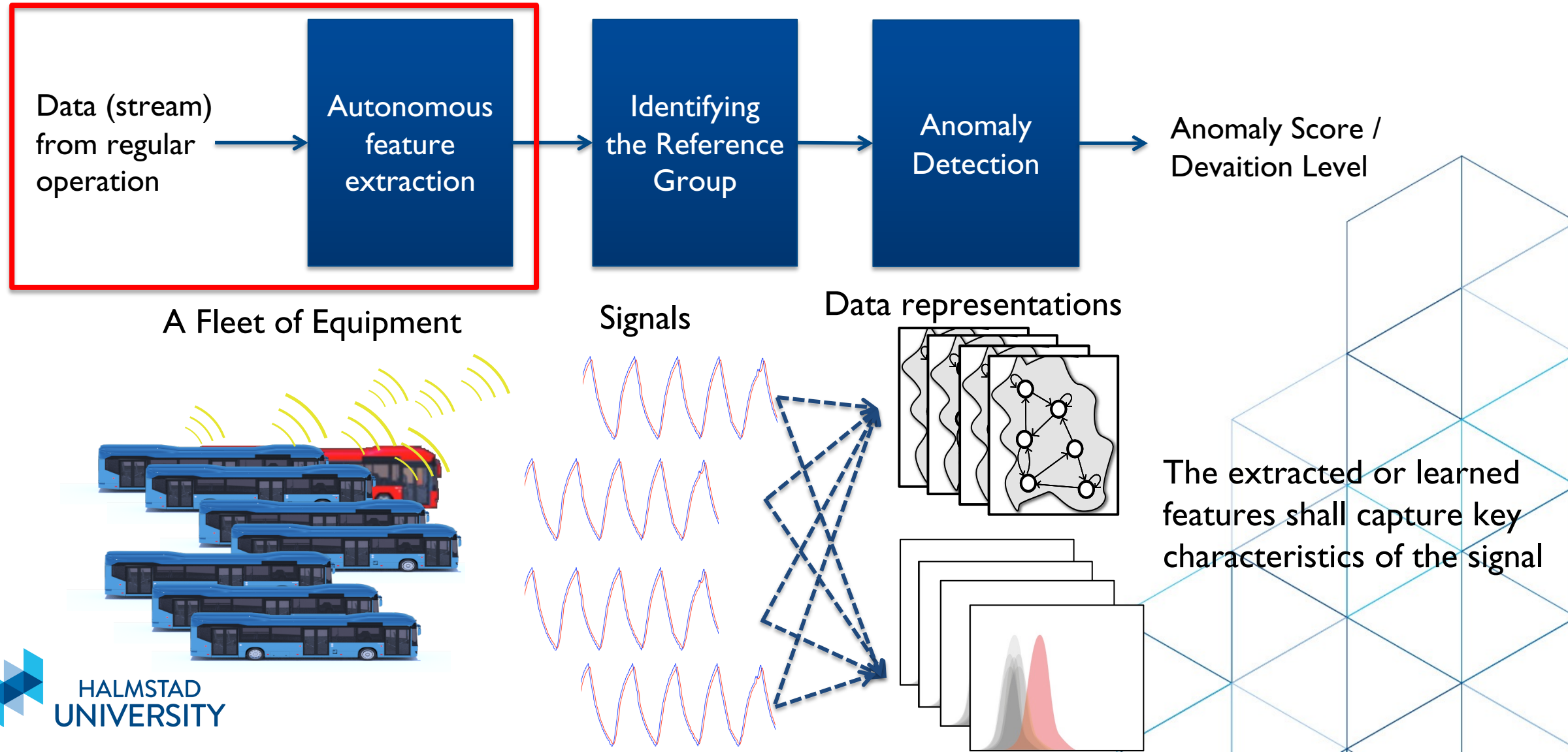
Lab I: Anomaly Detection with COSMO

Consensus Self-organizing Models (COSMO)

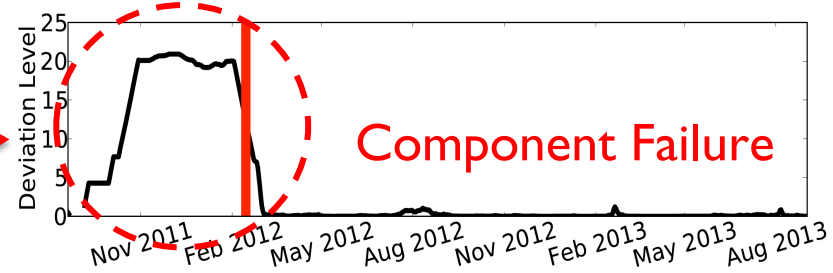
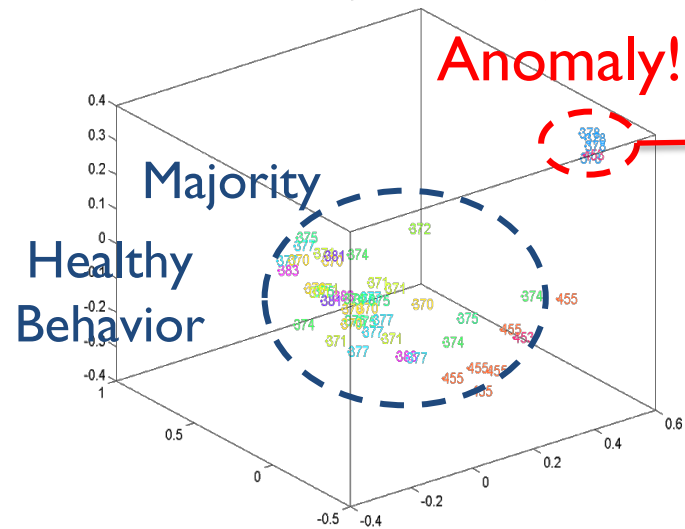
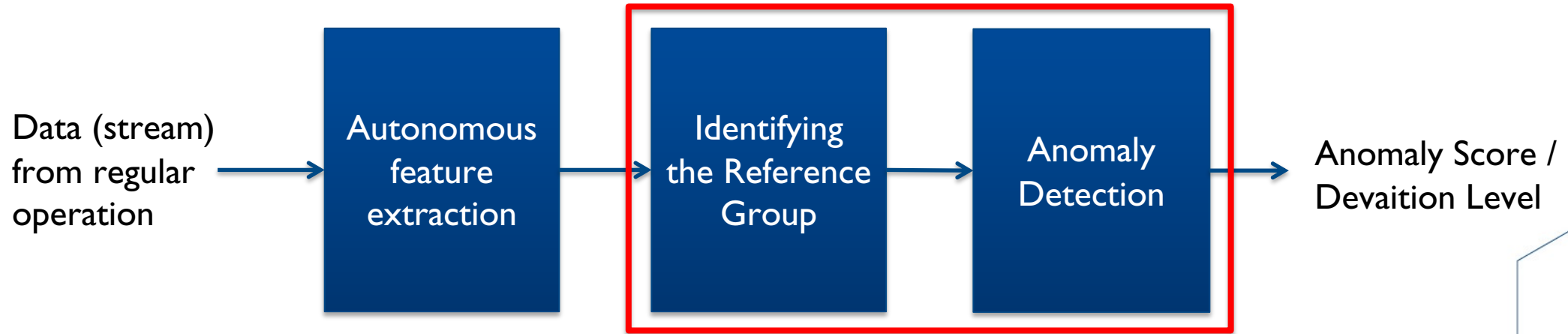
- Unsupervised anomaly detection
- “Wisdom of the Crowd”
 - Assume majority of the samples are healthy, deviations from majority is labelled as faulty
 - Feature extraction
 - Conformal anomaly detection



Consensus Self-organizing Models (COSMO)



Consensus Self-organizing Models (COSMO)



Fleet Operator

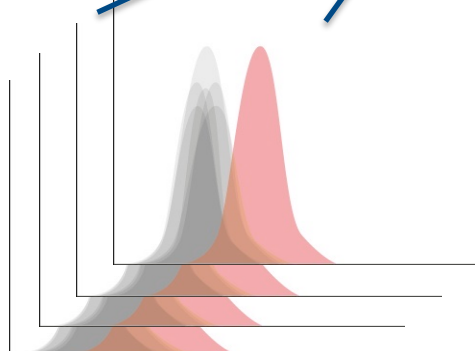


COSMO

Hellinger Distance

$$H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^k (\sqrt{P_i} - \sqrt{Q_i})^2}$$

Histogram



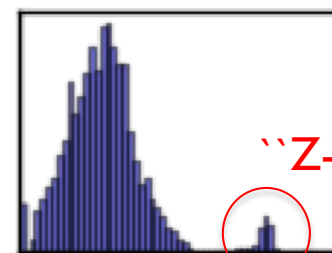
Matrix of pairwise distance

$$D = \begin{pmatrix} H_{v_1 d_1, v_1 d_1} & H_{v_1 d_2, v_1 d_1} & H_{v_1 d_3, v_1 d_1} & \cdots & H_{v_1 d_k, v_1 d_k} \\ H_{v_2 d_1, v_1 d_1} & H_{v_2 d_2, v_1 d_1} & H_{v_2 d_3, v_1 d_1} & \cdots & H_{v_2 d_k, v_1 d_k} \\ H_{v_3 d_1, v_1 d_1} & H_{v_3 d_2, v_1 d_1} & H_{v_3 d_3, v_1 d_1} & \cdots & H_{v_3 d_k, v_1 d_k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{v_k d_1, v_1 d_1} & \cdots & \cdots & \cdots & H_{v_k d_k, v_1 d_k} \end{pmatrix}$$

Most Central Pattern c

"Z-score"

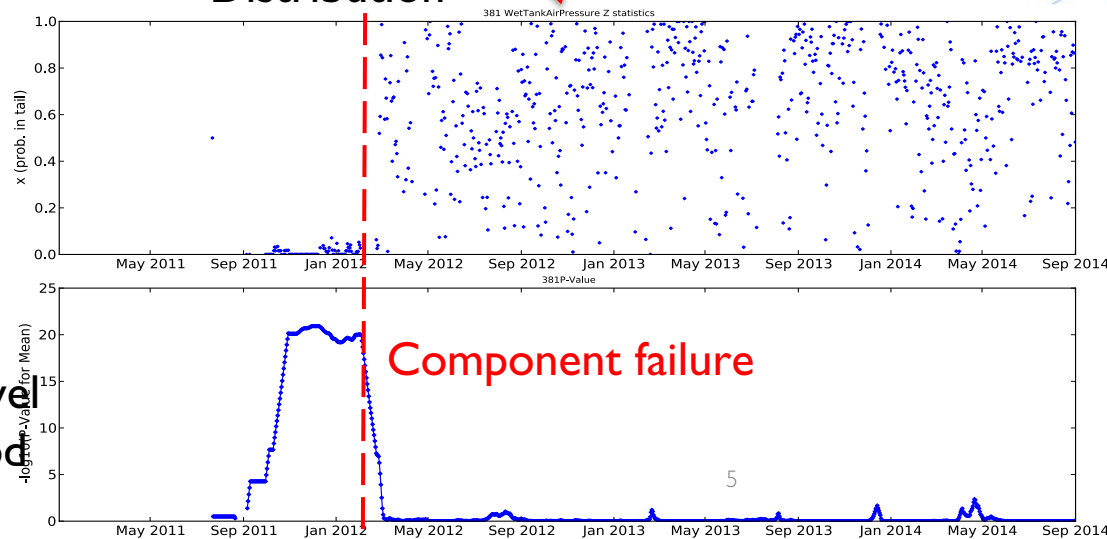
$$z(m) = \frac{|\{i = 1, \dots, N : d_{i,c} > d_{m,c}\}|}{N}$$



Empirical Distribution

Z-score

Deviation level over a period



Component failure

Anomaly Detection with COSMO

Algorithm 1: Compute anomaly score with COSMO method

Data: Time Series Data X ; /* $x_t^{i,j}$ for unit i feature set j at time t */
Time Length T ;
Reference Group R_t^i ;
Feature extraction method $g(\cdot)$;
Metric comparing two repr. $\delta(\cdot, \cdot)$;
Time length for uniformity test T_u

Result: Anomaly score ϵ_t^i

initialization;
Time Step $t \leftarrow 1$;
while $t \leq T$ **do**
 $\theta_t^i \leftarrow g(X_t^{i,j})$; /* Compute compressed representation/model) */
 Acquire reference group $R_t^i \leftarrow R_t^i \cup \theta_t^i$;
 Compute pairwise distance matrix \mathbf{D} based on $\delta(\cdot, \cdot)$
 Acquire most central pattern c and its row of distances \mathbf{d} to the rest in R_t^i ;
 Compute z-score $z(m) = \frac{|\{k=1,\dots,N:d_{k,c}>d_{m,c}\}|}{N}$; /* $m \leftarrow x_t^i$, $N = |R_t^i|$ */
 $t \leftarrow t + 1$;
end
 $t \leftarrow T_u$;
while $t \leq T$ **do**
 $\bar{z}_t^i \leftarrow \frac{1}{T_u} \sum_{t-T_u}^t z_t^i$;
 Compute Anomaly Score $\epsilon(\bar{z}) = -\log_{10} \left[\Phi \left(\frac{\bar{z}-0.5}{\sigma_n} \right) \right]$; /* $\sigma_n = (12n)^{-\frac{1}{2}}$ */
 $t \leftarrow t + 1$;
end

