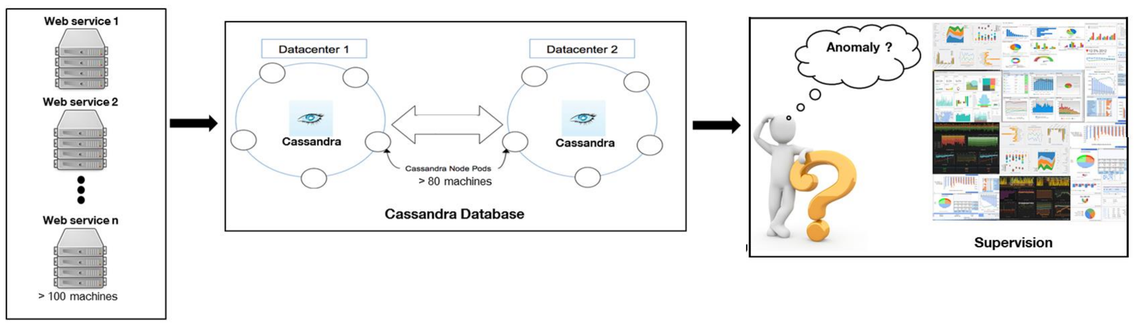
# Unsupervised Anomaly Detection for Cyber-Physical Systems

Audibert Julien
julien.audibert@orange.com
Orange / Eurecom

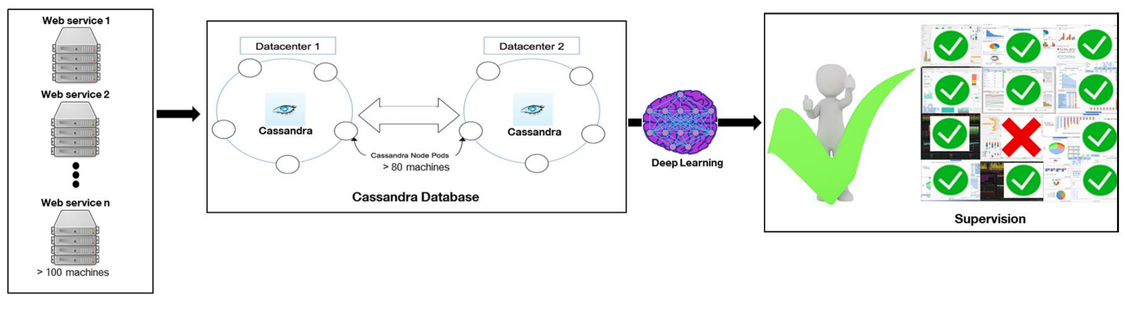
## Motivation




Challenges :

- Threshold-based anomaly detection does not capture dynamics of anomalies
- Lack of labeled data for supervised machine learning

## Objective




Goal :

- Detect anomalies in a real-world dynamic system such as Cyber Physical Systems*
- Explain decision making to supervisors

*CPS are integration of computation, networking and physical process

## Dataset

| Dataset | Open data | Days | Size | Number of logs | Environment |
|---|---|---|---|---|---|
| **Secure Water Treatment (SWaT)** | Yes | 11 | 335 MB | ~950 000 | Testbed |
| **Orange internal** | No | 90 | 40 TB | ~82 Billion | Production |

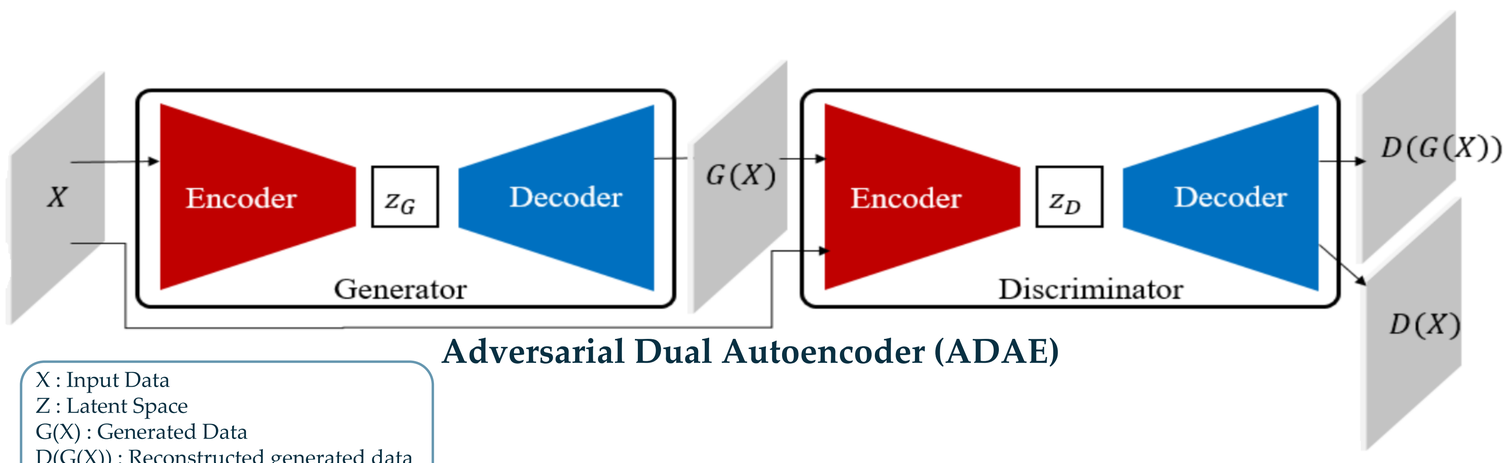
https://itrust.sutd.edu.sg/itrust-labs_datasets/dataset_info/

## Methodology

- Architecture of ADAE [2] with two autoencoders as generator and discriminator.
- Training objective :
  - Discriminator $\mathcal{L}_D = \|X - D(X)\|_2 - \|G(X) - D(G(X))\|_2$
  - Generator $\mathcal{L}_G = \|X - G(X)\|_2 + \|G(X) - D(G(X))\|_2$
- Anomaly score :

$$\mathcal{A}(\hat{x}) = \|\hat{x} - \mathrm{D}(\mathrm{G}(\hat{x}))\|_2$$

## Architecture




**Adversarial Dual Autoencoder (ADAE)**

X : Input Data
Z : Latent Space
G(X) : Generated Data
D(G(X)) : Reconstructed generated data
D(X) : Reconstructed X

## Preliminary Results

SWaT Dataset

| Methods | Precision | Recall | F1 score |
|---|---|---|---|
| PCA | 24.92 | 21.63 | 0.23 |
| KNN | 7.83 | 7.83 | 0.08 |
| AE | 72.63 | 52.63 | 0.61 |
| MAD-GAN [3] | 98.97 | **63.74** | **0.77** |
| **ADAE** | **99.38** | 62.19 | **0.77** |

## Ongoing Research

- Stabilize GAN training
- Try Gated Recurrent Units to improve performance
- Test on Orange internal Dataset
- Add interpretability layer

## Conclusions

- **High potential** to detect anomalies on complex dynamic systems such as Cyber Physical Systems.
- GAN-based architecture makes **learning complicated**.

## References

[1] Chalapathy, R., & Chawla, S. (2019). *Deep Learning for Anomaly Detection: A Survey.*
[2] Vu, H. S., Ueta, D., Hashimoto, K., Maeno, K., Pranata, S., & Shen, S. M. (2019). *Anomaly Detection with Adversarial Dual Autoencoders.*
[3] Li, D., Chen, D., Shi, L., Jin, B., Goh, J., & Ng, S.-K. (2019). *MAD-GAN: Multivariate Anomaly Detection for Time Series Data with Generative Adversarial Networks.*
[4] Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., & Cottrell, G. W. (n.d.). *A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction.*

## Contact

Audibert Julien
Orange / Eurecom
julien.audibert@orange.com

DS3 DATA SCIENCE SUMMER SCHOOL
Monday, June 24 – Friday, June 28, 2019

ÉCOLE POLYTECHNIQUE
UNIVERSITÉ PARIS-SACLAY