Séminaire interne

29/09/2023
Côme Frappé - - Vialatoux

Sujet de thèse

Détection d'anomalie dans les Réseaux de distribution de l'eau par algorithmes génétiques à estimation de distribution

Étude de cas - Données cyber-physiques

Plan :

1. Exemple de données physique
2. Exemple d'interactions cyber-physiques
3. Résultats récents : Méthodologie d'exploration de données Cyber

Données physiques

Description du jeu de données:

Nom : WADI (lien)

- ~ 1M lignes
- 90+ colonnes
  - Senseurs physiques
    - Flux, niveau, valves, pression, pompes, etc.
  - Senseurs chimiques
    - Ph, turbidité, conductivité, potentiel redox, chlorine résiduelle
- 13 Scénarios d'attaques

Exemple

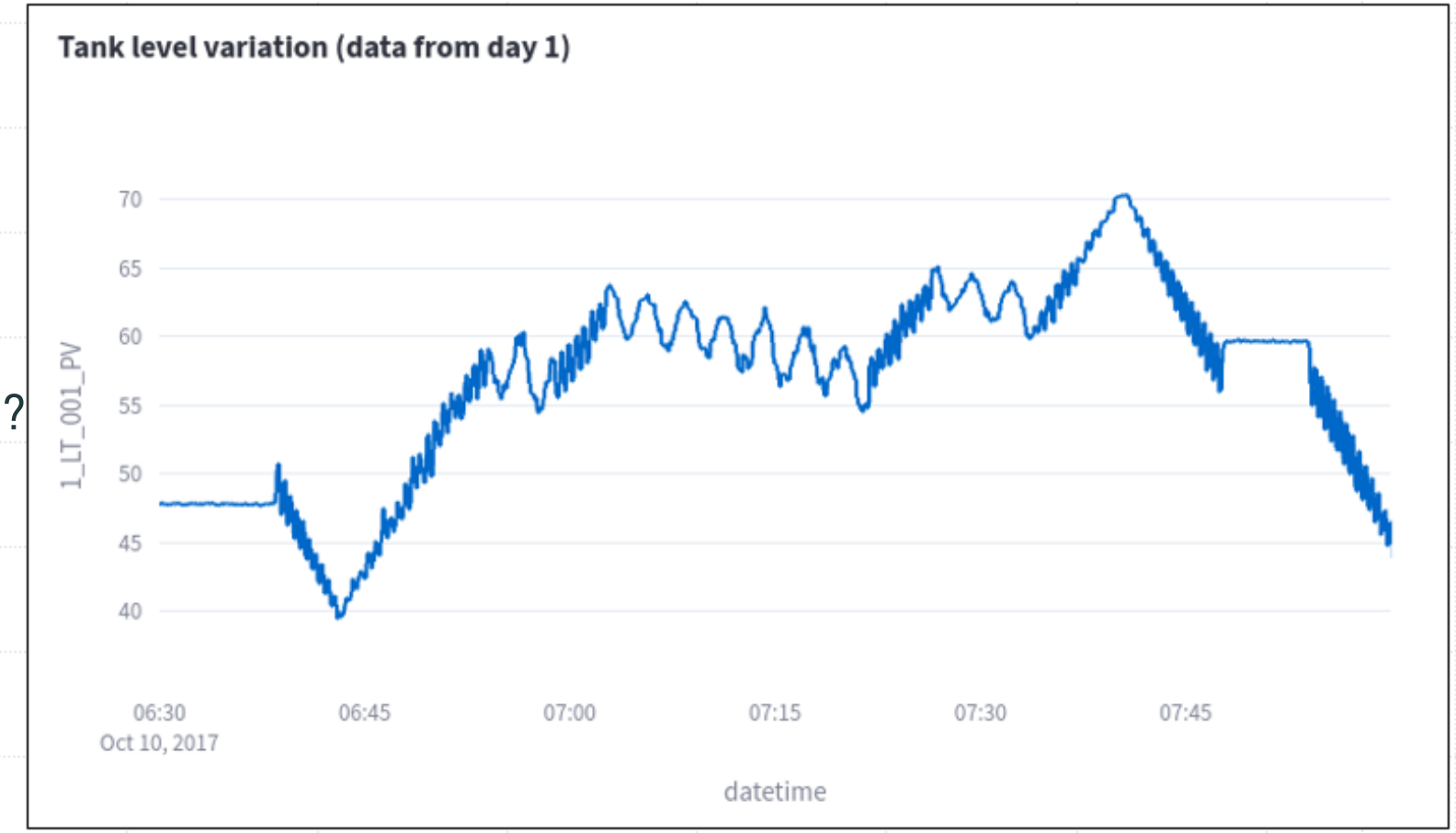
Attack Identifier	Starting Time	Ending Time	Duration (minutes)	Attack description
1	9/10/17 19:25:00	9/10/17 19:50:16	25.16	Motorized valve 1_MV_001 is maliciously turned on, this causes an overflow on primary tank **should reflect on 1LT001 and 1FIT001**

Données physiques

Parmi les Données:

Niveau d'eau dans une cuve au cours du temps

Question : Que remarquez-vous ?



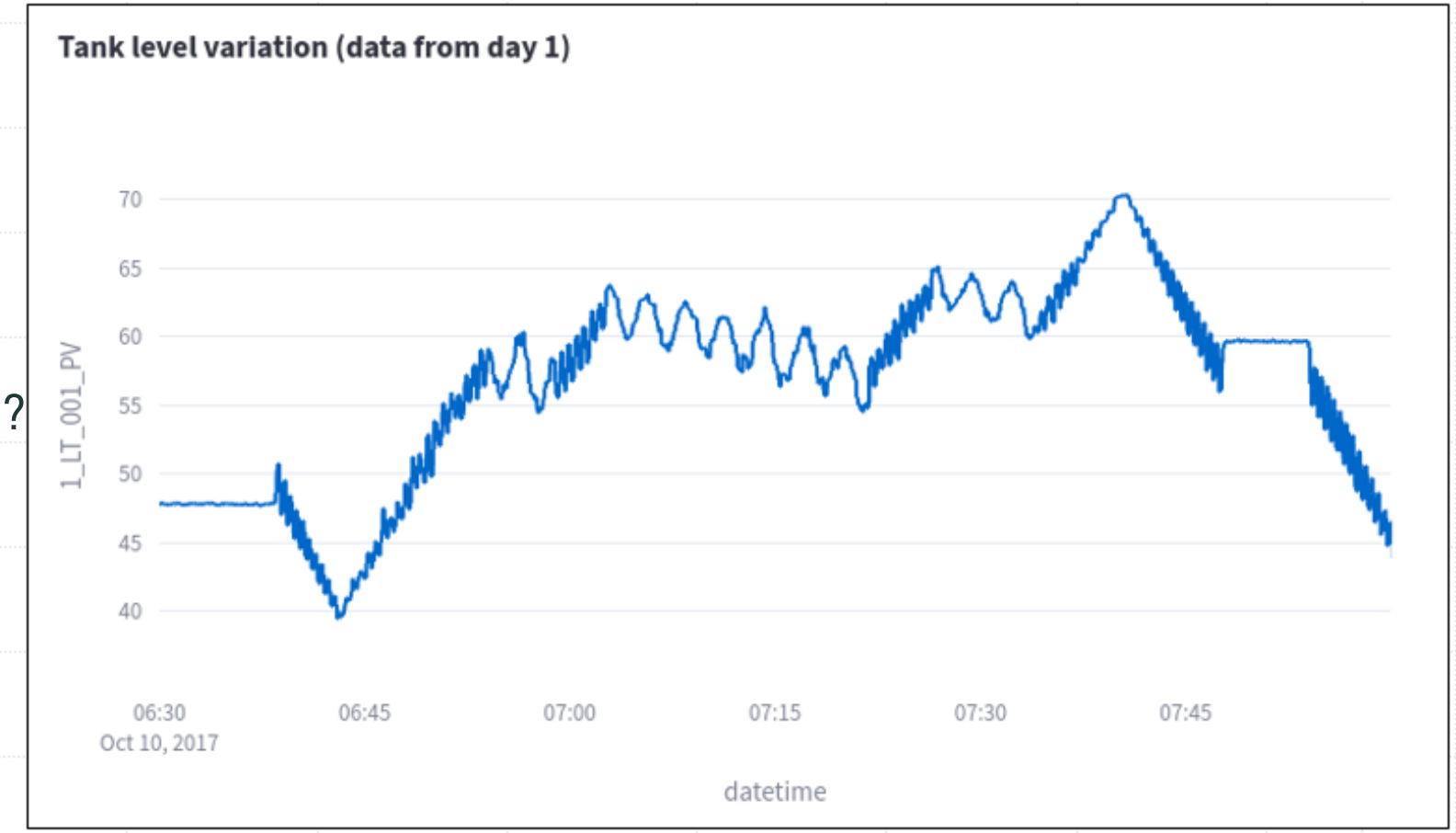
Données physiques

Parmi les Données:

Niveau d'eau dans une cuve au cours du temps

Question : Que remarquez-vous ?

- passage d'un palier à l'autre en "Oscillation"

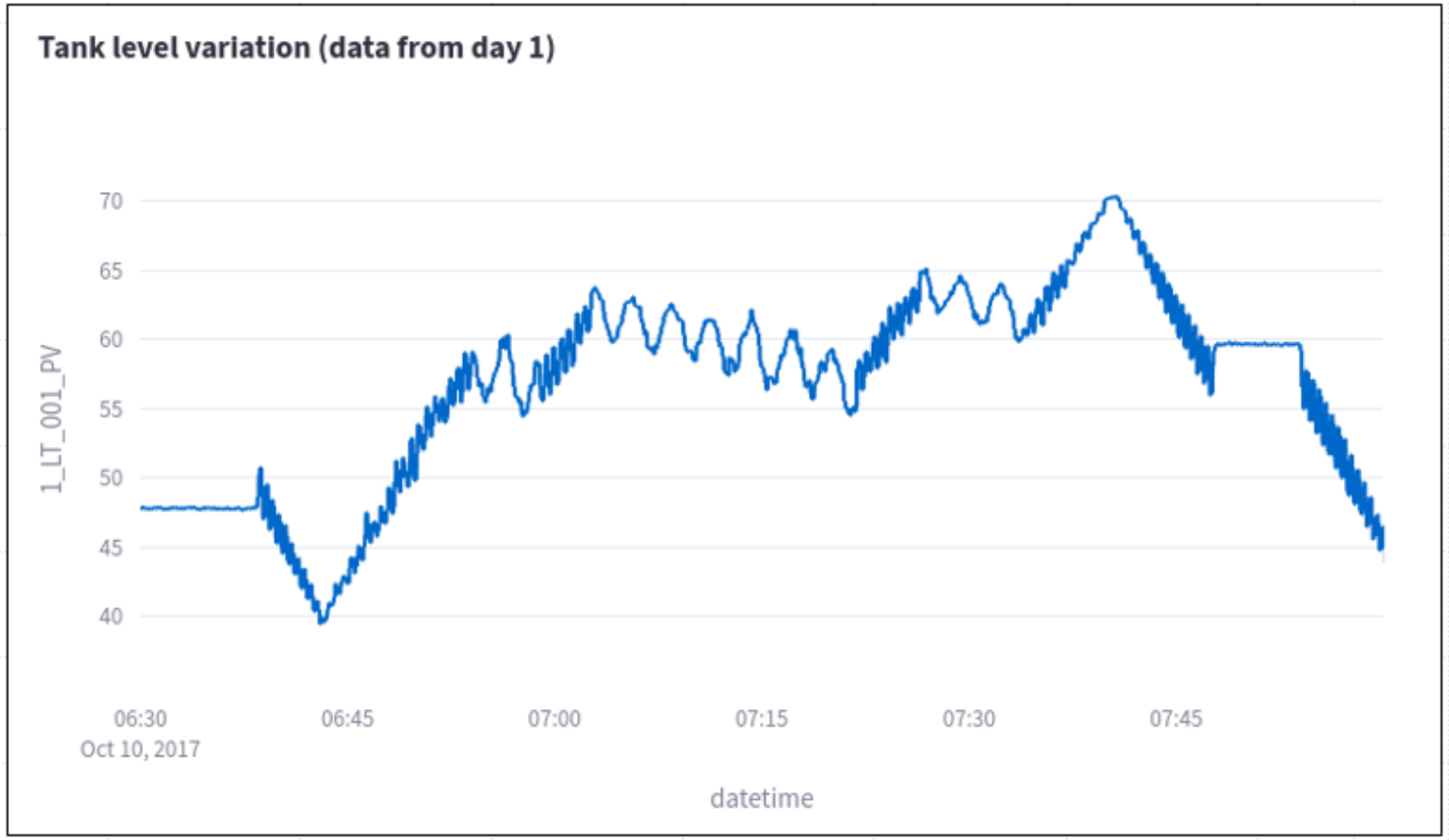


Données physiques

Parmi les Données:

Niveau d'eau dans une cuve au cours du temps

- Passage d'un palier à l'autre en "Oscillation"
  -> Explication ?

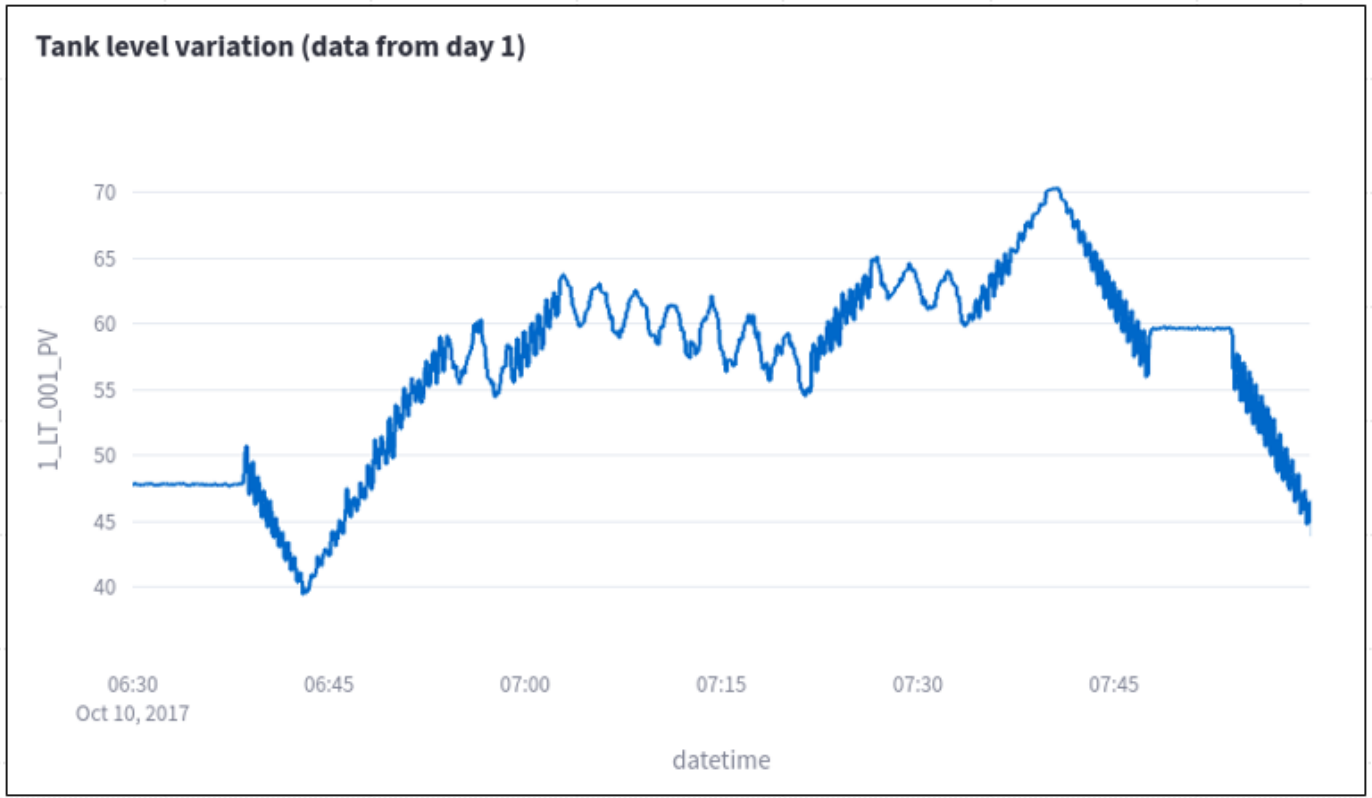


Données physiques

Parmi les Données:

Niveau d'eau dans une cuve au cours du temps

- Passage d'un palier à l'autre en "Oscillation"
- Comment mesurer le niveau d'eau ?

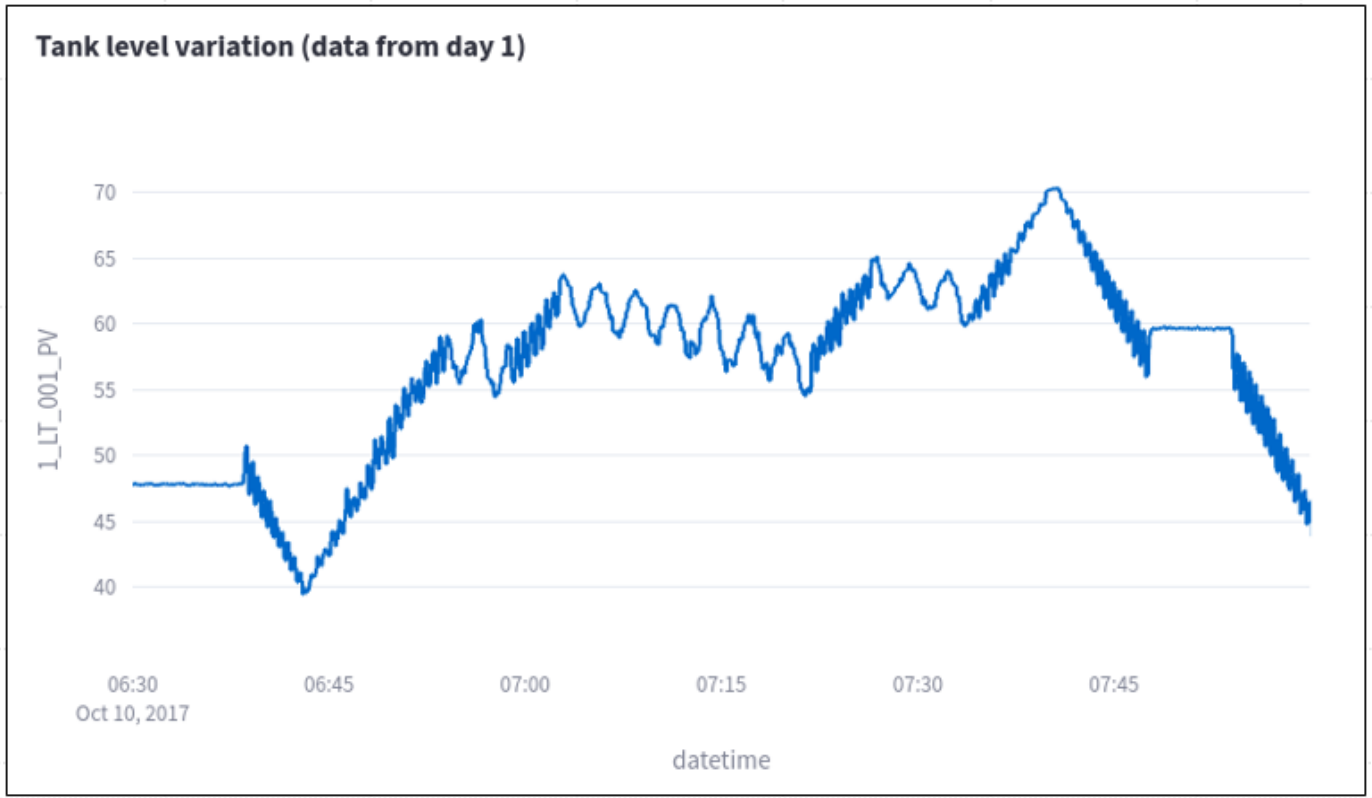


Données physiques

Parmi les Données:

Niveau d'eau dans une cuve au cours du temps

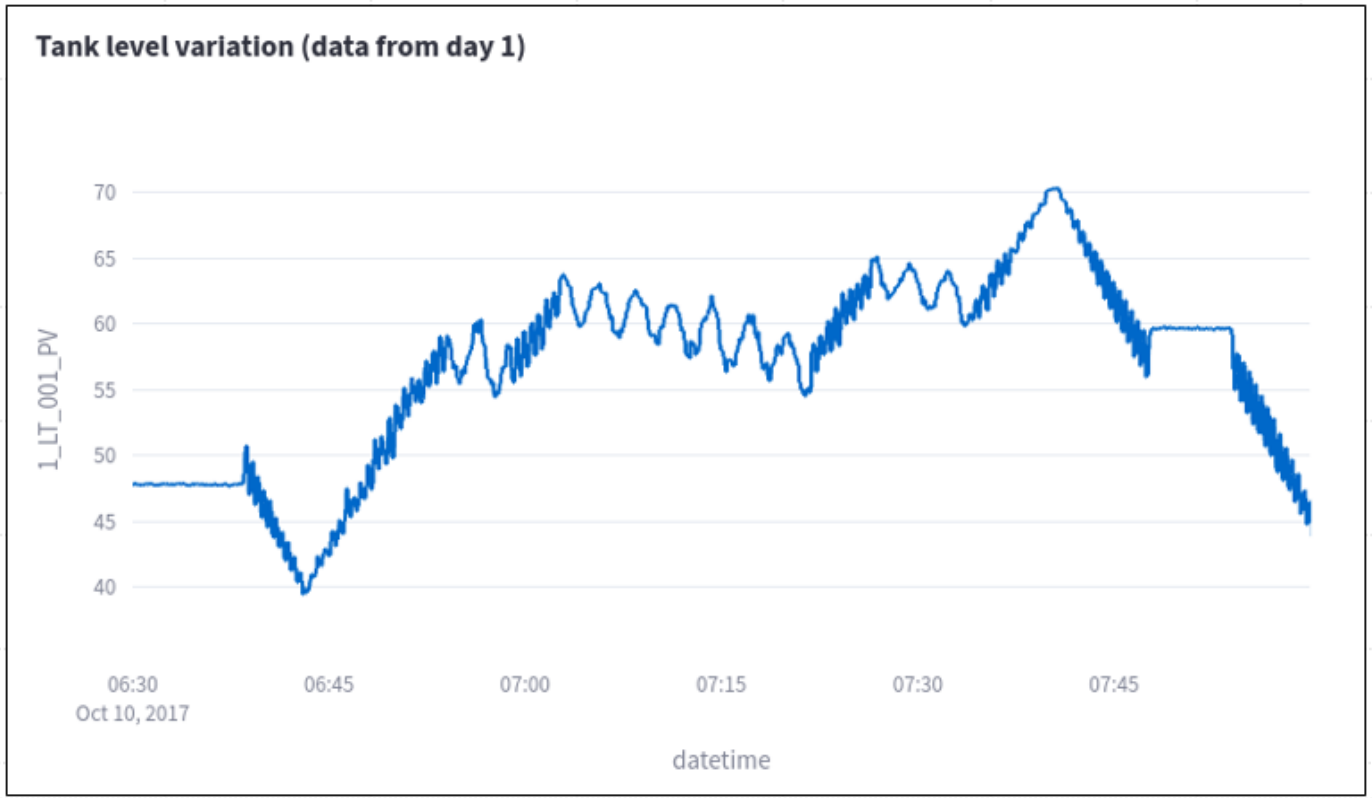
- Passage d'un palier à l'autre en "Oscillation"
- Comment mesurer le niveau d'eau ?
  - Capteur de Surface
  - Capteur de pression



Données physiques

Conclusion :

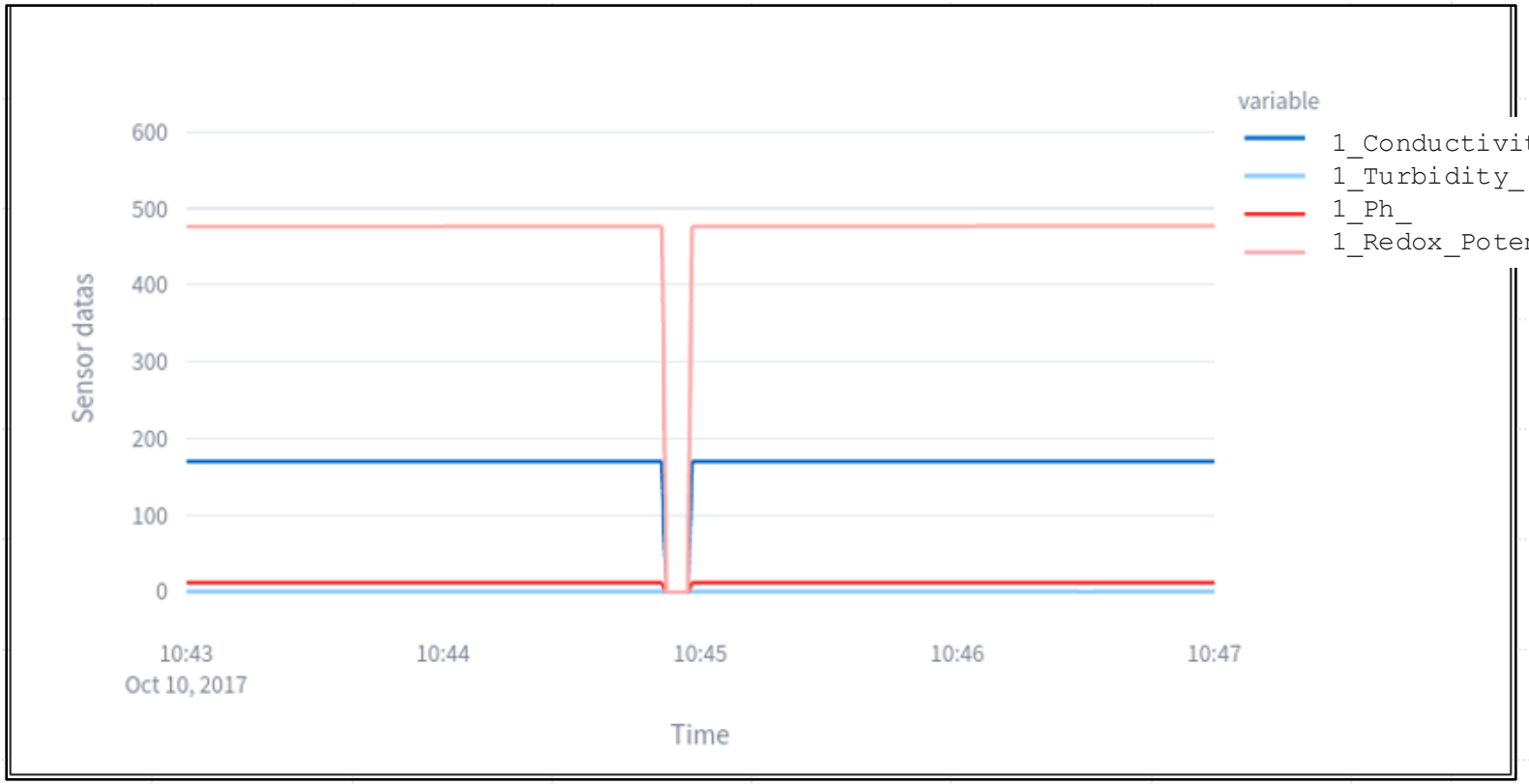
- Visualisation du niveau d'eau => inférer le type d'équipement de capture



Données physiques

Autre exemple:

- Pendant 6 secondes, les valeurs de 4 colonnes passent à zéro
- Que s'est-il passé ?



Données physiques

Autre exemple:

- Pendant 6 secondes, les valeurs de 4 colonnes passent à zéro
- Que s'est-il passé ?

```
1_Conductivity_
1_Turbidity_
1_Ph_
1_Redox_Potential
```

Table 2: Equipment list for P1 (Water Supply)

Description	Design Specification	Material	Qty	Brand & Model	Remarks
Pumps & Tanks					
Raw Water Tank	Capacity: 2.5m³	Tank Wall: FRP Stand & Base: Mild Steel	2	Customised	With Drain Valve at Bottom
Raw Water Transfer Pump	Duty: 2.5 m³/h @ 20m	Casing: Chrome Nickel SS Impeller: Noryl Shaft: SS	2	CALPEDA MXH 203	
Chemical Dosing Pumps	Capacity : 0.78 l/h @ 5 bar	Liquid end : Plexiglas Diaphragm : PTFE faced	4	Prominent GALa1601	
Instrumentation					
Level Transmitters	Radar, Range 0.2 to 6m	Non Contact	1	iSOLV RD 700	
Flow Transmitter	Electromagnetic DN40	PTFE	1	iSOLV EFS803/CFT183	
Multi Probe Analyser	pH/ ORP/ Conductivity & Turbidity		1	Hydrolab HL4	
Total Chlorine Analyser	TRC 0-5ppm		1	W&T Depolox 3	
Piping & Accessories					
Piping	SCH80	PVC	Lot		
On/Off Valve	DN 25, Electric Actuated	PVC	3	Burkert Type 3003	

Données physiques

Autre exemple:

- Pendant 6 secondes, les valeurs de 4 colonnes passent à zéro
- Que s'est-il passé ?

```
1_Conductivity_
1_Turbidity_
1_Ph_
1_Redox_Potential
```

Table 2: Equipment list for P1 (Water Supply)

Description	Design Specification	Material	Qty	Brand & Model	Remarks
Pumps & Tanks					
Raw Water Tank	Capacity: 2.5m³	Tank Wall: FRP Stand & Base: Mild Steel	2	Customised	With Drain Valve at Bottom
Raw Water Transfer Pump	Duty: 2.5 m³/h @ 20m	Casing: Chrome Nickel SS Impeller: Noryl Shaft: SS	2	CALPEDA MXH 203	
Chemical Dosing Pumps	Capacity : 0.78 l/h @ 5 bar	Liquid end : Plexiglas Diaphragm : PTFE faced	4	Prominent GALa1601	
Instrumentation					
Level Transmitters	Radar, Range 0.2 to 6m	Non Contact	1	iSOLV RD 700	
Flow Transmitter	Electromagnetic DN40	PTFE	1	iSOLV EFS803/CFT183	
Multi Probe Analyser	pH/ ORP/ Conductivity & Turbidity		1	Hydrolab HL4	
Total Chlorine Analyser	TRC 0-5ppm		1	W&T Depolox 3	
Piping & Accessories					
Piping	SCH80	PVC	Lot		
On/Off Valve	DN 25, Electric Actuated	PVC	3	Burkert Type 3003	

Données physiques

Autre exemple:

- 4 données indépendantes
- Mais captées par la même sonde
- Défaillance technique
  - 4 données impactées en même temps

Table 2: Equipment list for P1 (Water Supply)

Description	Design Specification	Material	Qty	Brand & Model	Remarks
Pumps & Tanks					
Raw Water Tank	Capacity: 2.5m³	Tank Wall: FRP Stand & Base: Mild Steel	2	Customised	With Drain Valve at Bottom
Raw Water Transfer Pump	Duty: 2.5 m³/h @ 20m	Casing: Chrome Nickel SS Impeller: Noryl Shaft: SS	2	CALPEDA MXH 203	
Chemical Dosing Pumps	Capacity : 0.78 l/h @ 5 bar	Liquid end : Plexiglas Diaphragm : PTFE faced	4	Prominent GALa1601	
Instrumentation					
Level Transmitters	Radar, Range 0.2 to 6m	Non Contact	1	iSOLV RD 700	
Flow Transmitter	Electromagnetic DN40	PTFE	1	iSOLV EFS803/CFT183	
Multi Probe Analyser	pH/ ORP/ Conductivity & Turbidity		1	Hydrolab HL4	
Total Chlorine Analyser	TRC 0-5ppm		1	W&T Depolox 3	
Piping & Accessories					
Piping	SCH80	PVC	Lot		
On/Off Valve	DN 25, Electric Actuated	PVC	3	Burkert Type 3003	

Données physiques

- Comprendre les données = extraire de l'information supplémentaire
- Comprendre l'infrastructure = comprendre les données

Interractions Cyber-physiques

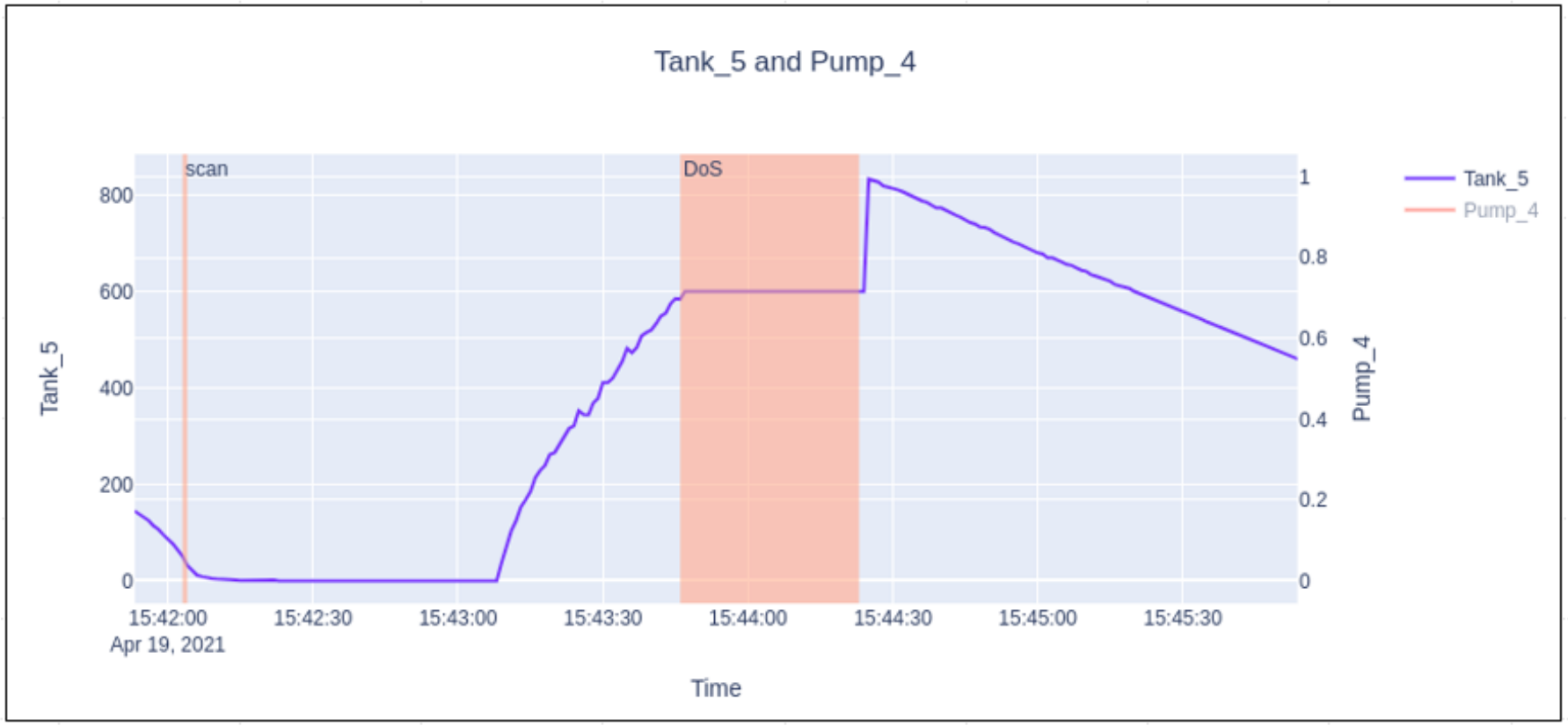
Description du jeu de données:

Nom : Hardware in the loop (Lien vers Dataset)

- Attaques physiques + cyber
  - Physical Fault, Man in the Middle, Dos, Scan
- Données Physiques:
  - ~11000 lignes
  - ~40 senseurs physiques (valves, tank level, pump flow pressure)
  - Pas de senseurs chimiques
- Données Réseau:
  - ~10M lignes
  - Communications des équipements

Interractions Cyber-physiques

Effet d'une attaque Cyber sur données physiques



Interractions Cyber-physiques

Effet d'une attaque Cyber sur données physiques

- Attaque doS sur un senseur = illusion de stabilité de niveau d'eau
- Caractérisation du DoS très forte grâce aux données physiques
  - Et si on faisait un pattern ?

Interractions Cyber-physiques

Effet d'une attaque Cyber sur données physiques

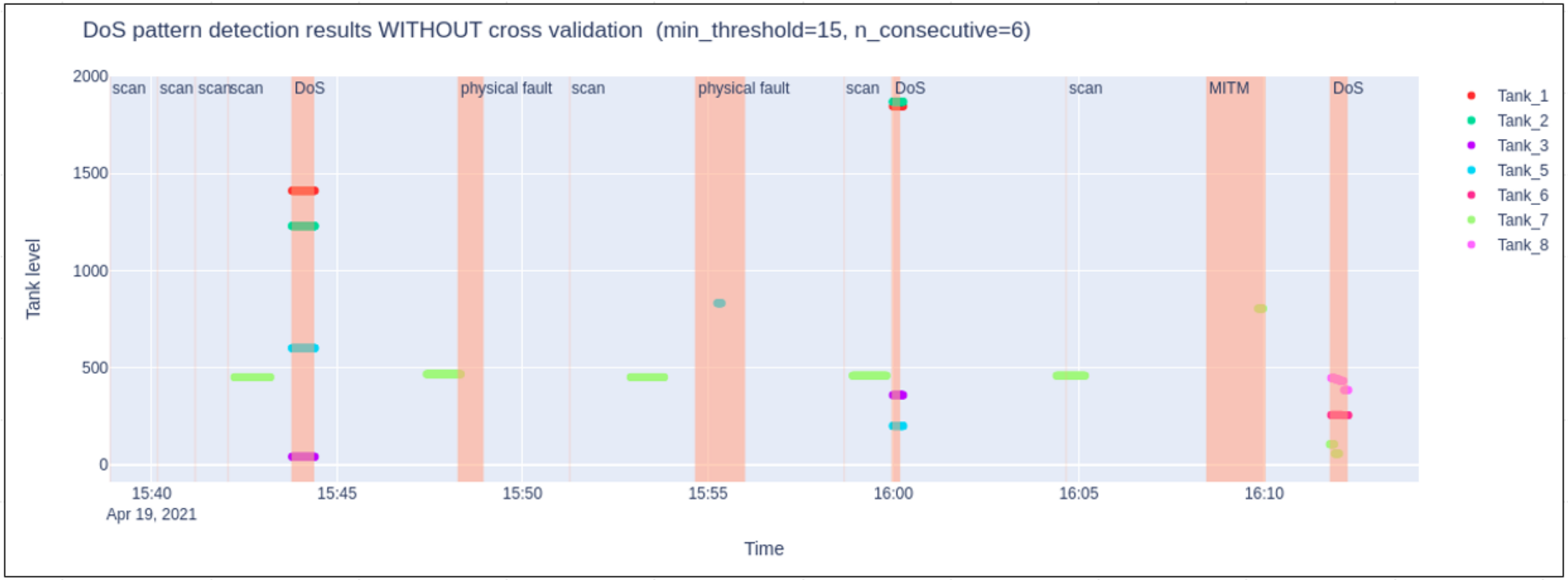
- Algorithme de détection de DoS par les données physiques
  - Pattern:

N_consecutive (constant values > Seuil_minimum) Sur N_tanks

Paramètres: ***N_consecutive*** : Durée Minimum de Dos à détecter
Seuil_minimum : Cuve vide ≠ DoS
N_tanks : nombre de tanks simultanés sur lequel pattern = Vrai
-> Réduit faux positifs

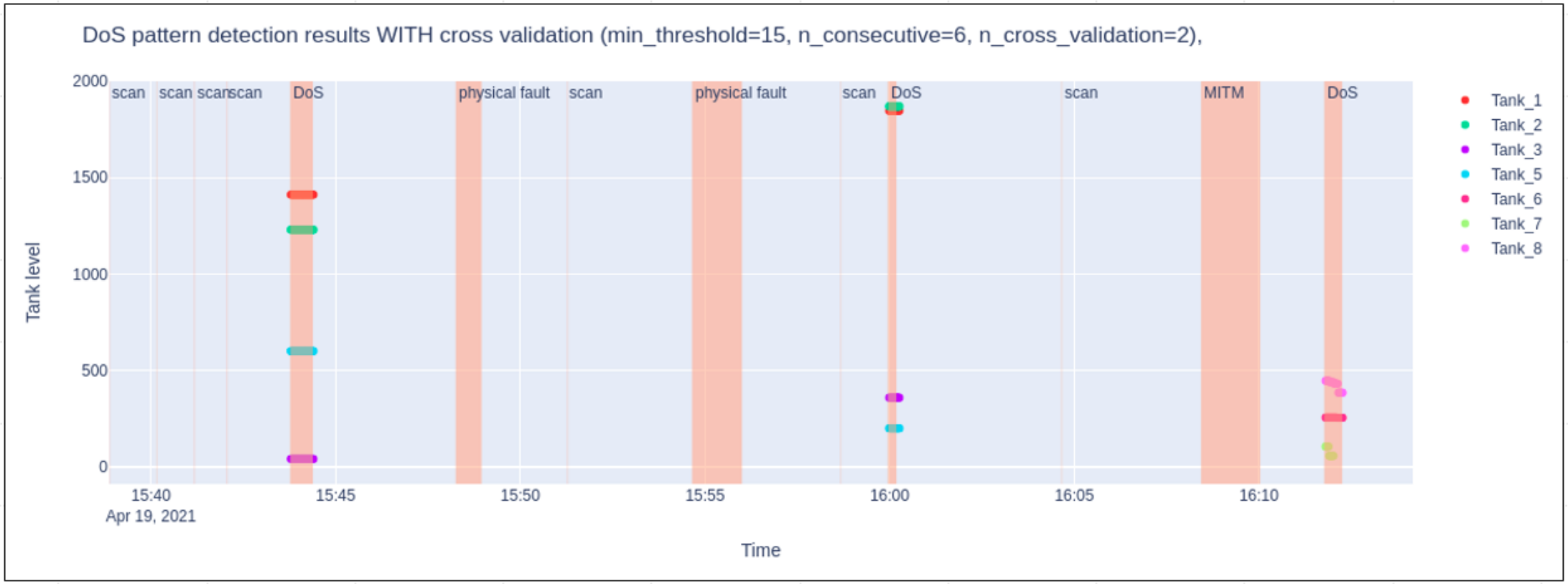
Interractions Cyber-physiques

Résultats (**SANS** utilisation du paramètre n_tank)



Interractions Cyber-physiques

Résultats (**AVEC** utilisation du paramètre n_tank=2)

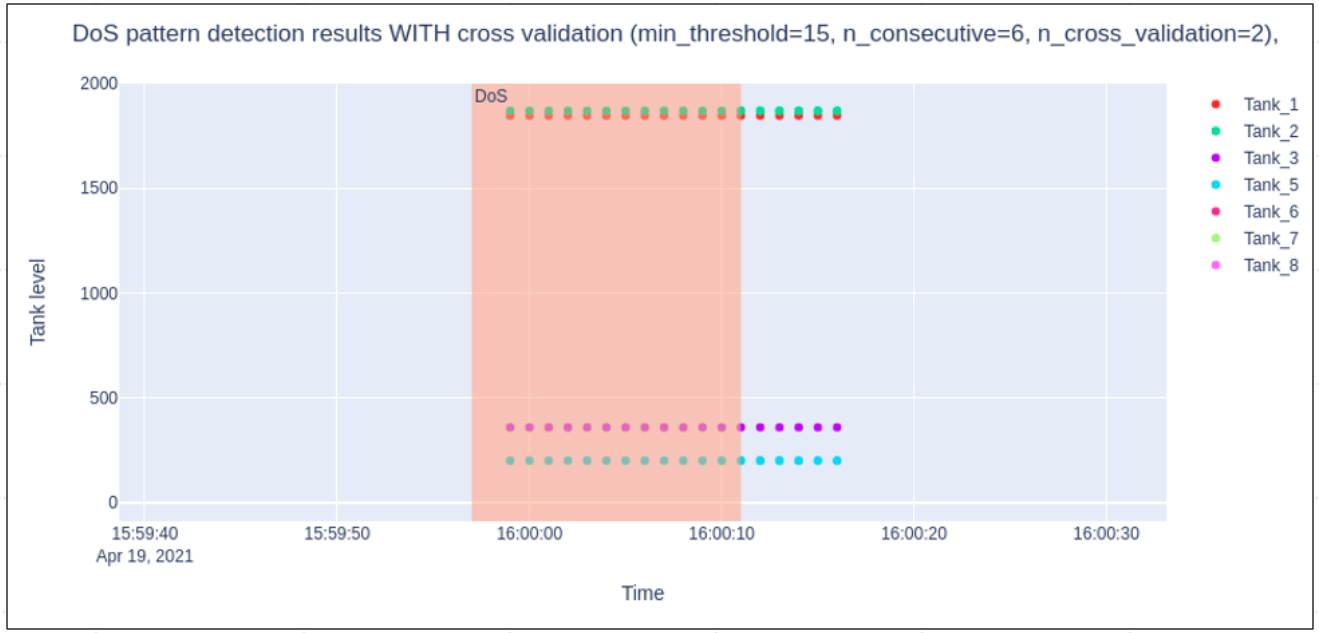


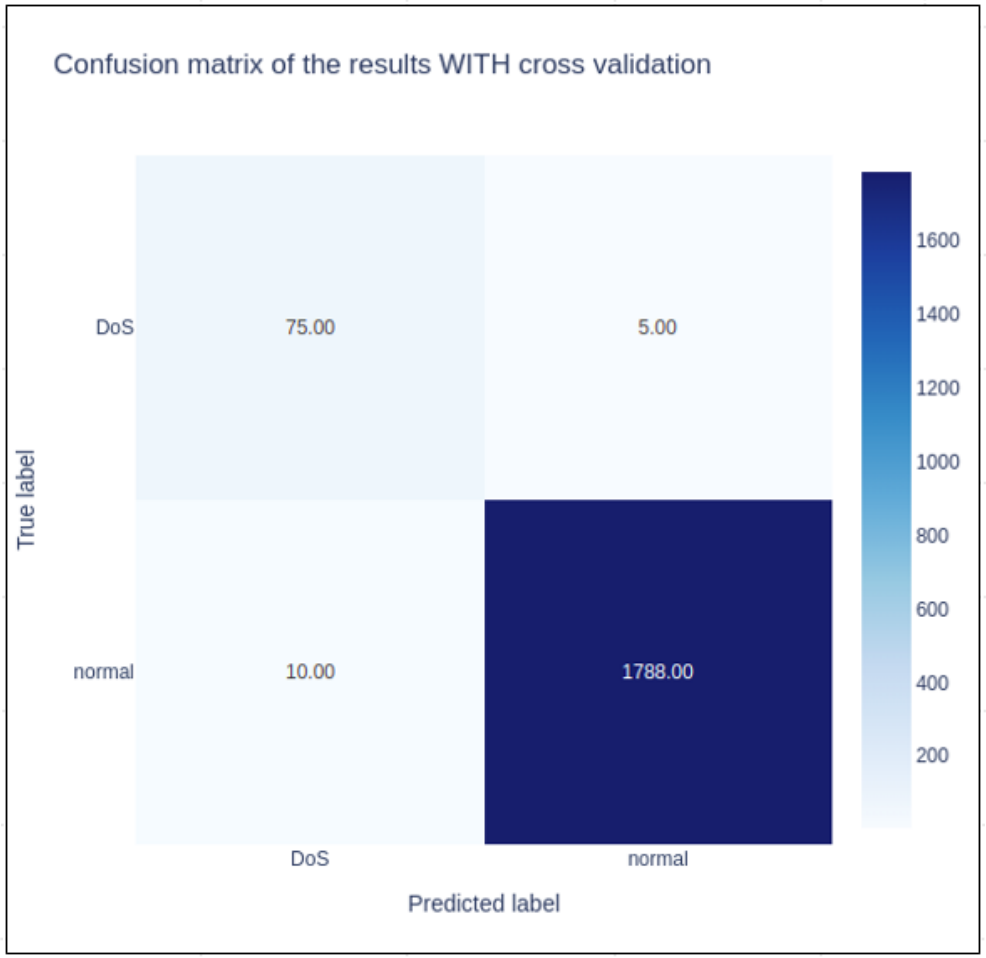
Interractions Cyber-physiques

Résultats (**AVEC** utilisation du paramètre n_tank=2)

FN = temps avant effet du DoS

FP = Temps après effet du DoS (reprise du système)





Interractions Cyber-physiques

Effet d'une attaque Cyber sur données physiques

- Effets interliés des attaques cyber <-> physiques
- Mise en relation des 2 plans = augmentation de surface détection

Résultats Récents

- Formalisation méthodologique
  - Analyse exploratoire de Datasets cyber-physiques

Approche Visuelle & hiérarchique :

- *Caractéristiques visualisation Haut-niveaux:*
  - Donnent une vue d'ensemble des données
  - Expliquent les concepts de base
  - Contiennent de l'information exclusive (pas présentes dans d'autres Viz)
  - Introduisent les Viz plus compliquées

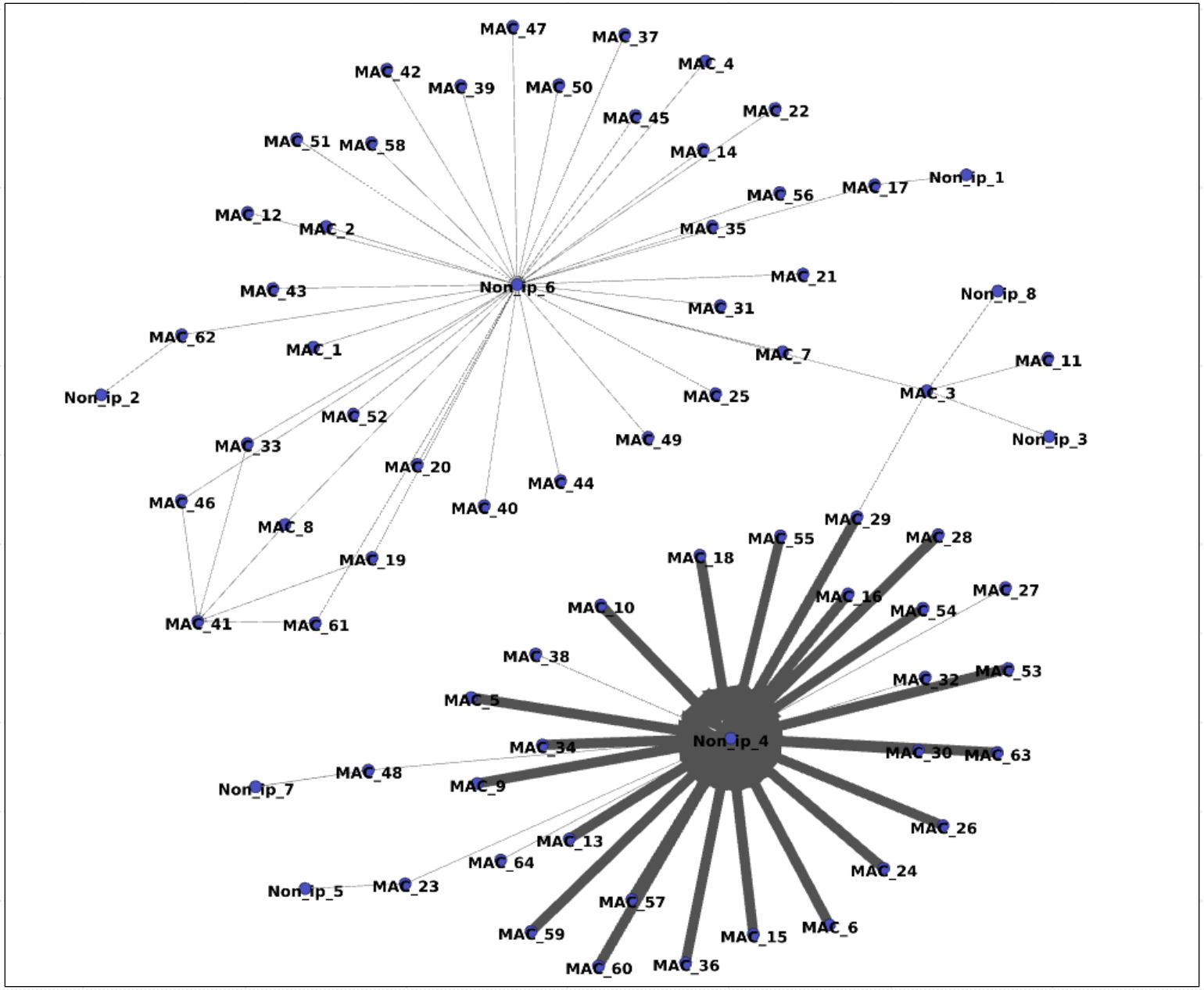
Résultats Récents

- Formalisation méthodologique
- 3 visualisations haut niveaux trouvées:
  - Topologique / Temporelle / Distribution
  - Méthodologie:
    1. Visualisation Topologique
       1. Identification de points d'intérêts
    2. Visualisation Temporelle
       1. Identification de points d'intérêts
    3. Visualisation de distribution
       1. Identification de points d'intérêts

Visualisation Topologique

Structure du réseau
- Qui parle avec qui
- Nombre de messages échangés

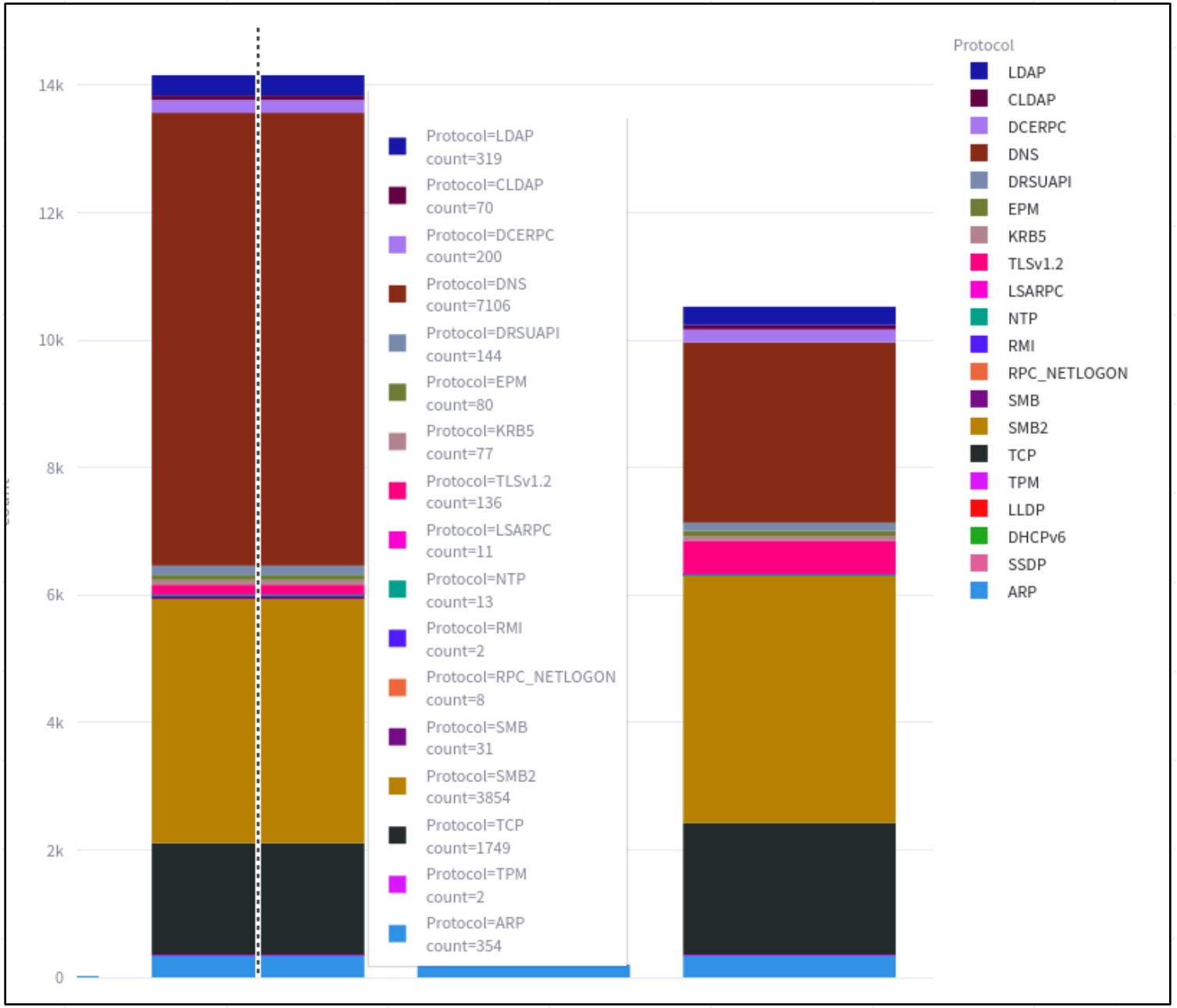
Vue d'ensemble → Check
Concept de base → Check
Information Exclusive → Check
Introduction d'autre Viz → Check



Visualisation Topologique

Structure du réseau
- Qui parle avec qui
- Nombre de messages échangés

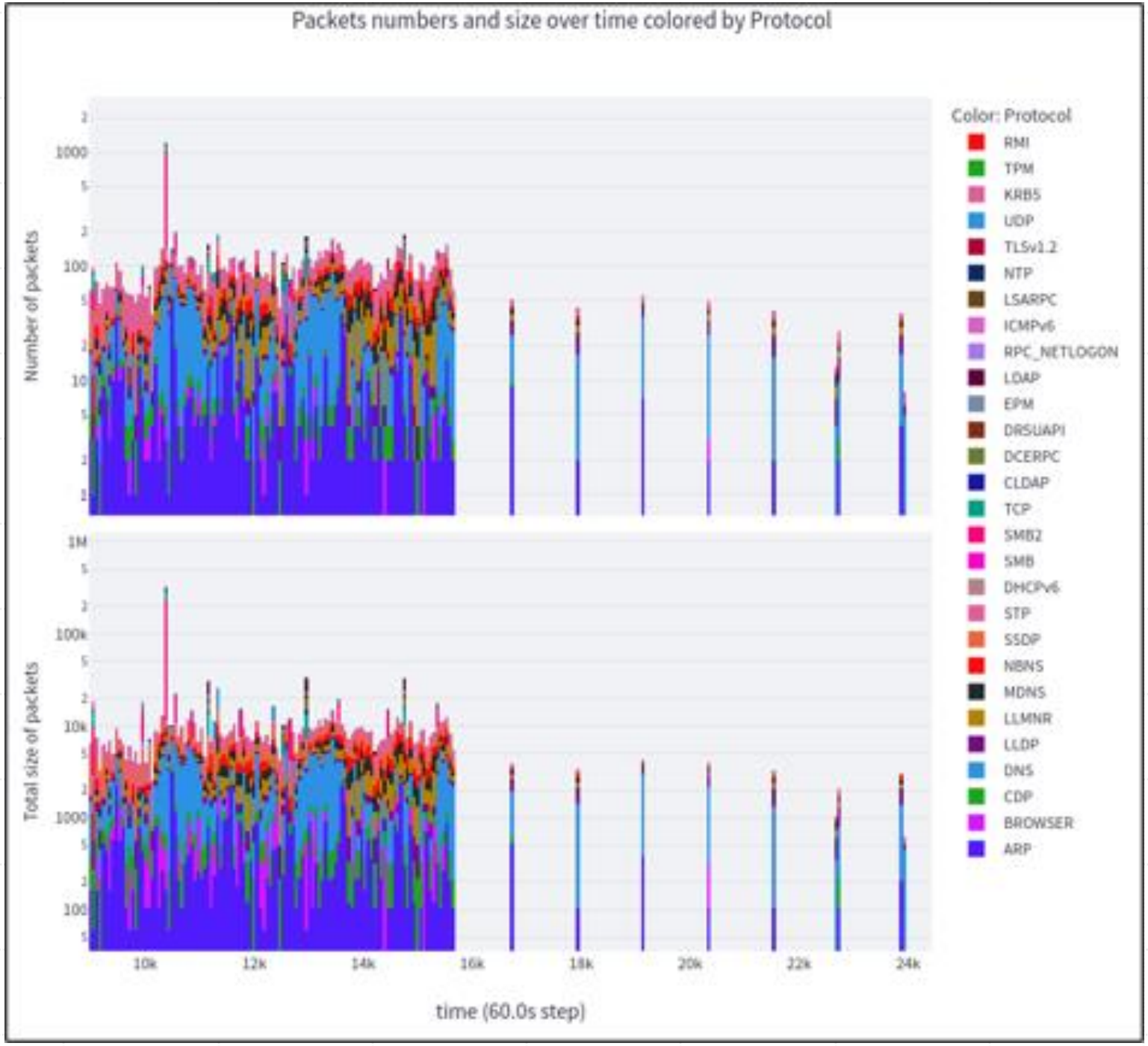
Vue d'ensemble → Check
Concept de base → Check
Information Exclusive → Check
Introduction d'autre Viz → Check



Visualisation Temporelle

- Protocoles utilisés dans le temps + nombre

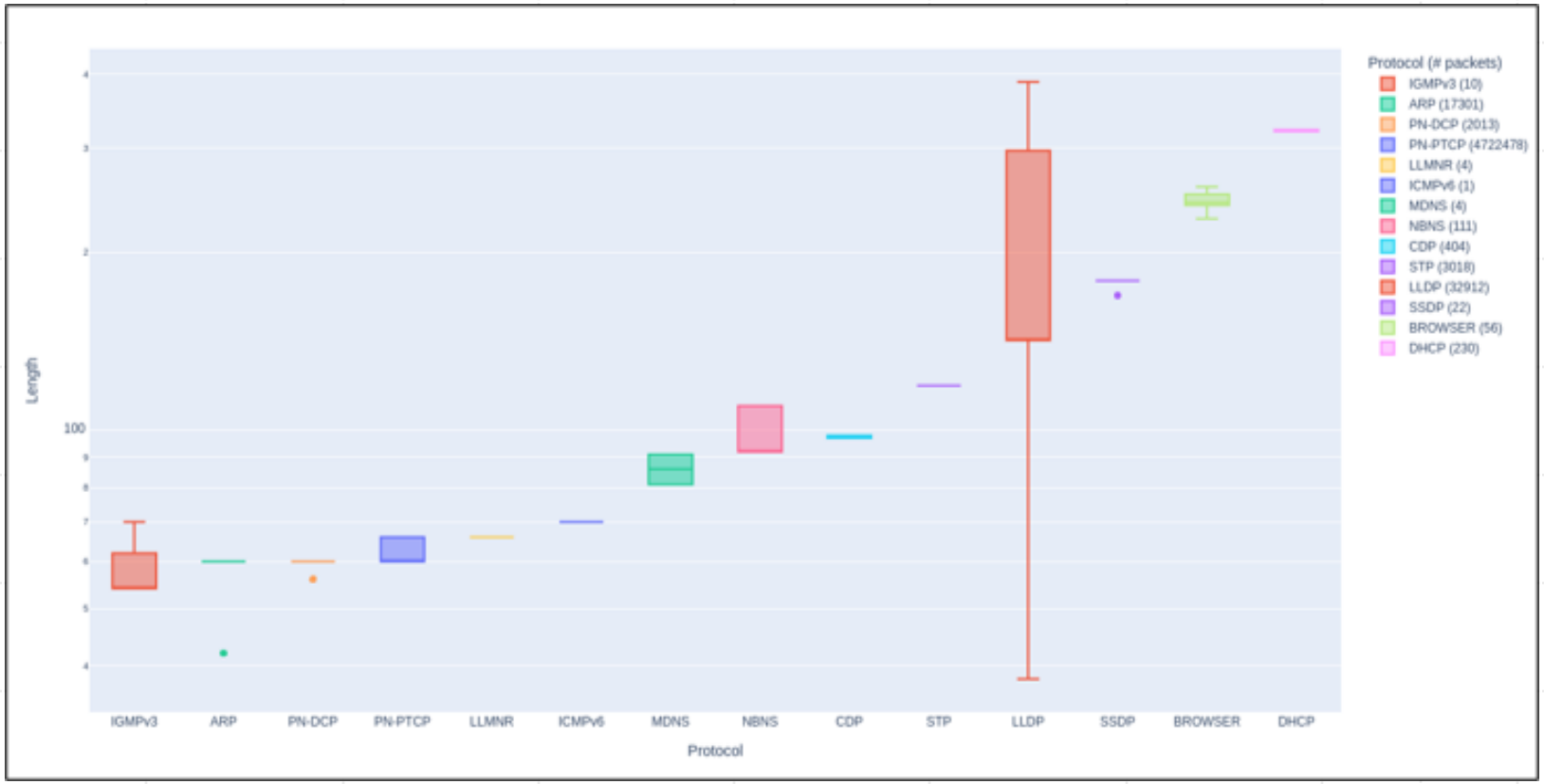
Vue d'ensemble → Check
Concept de base → Check
Information Exclusive → Check
Introduction d'autre Viz → ~Check
- Possibilité de naviguer



Visualisation Distribution

- Boxplot protocoles

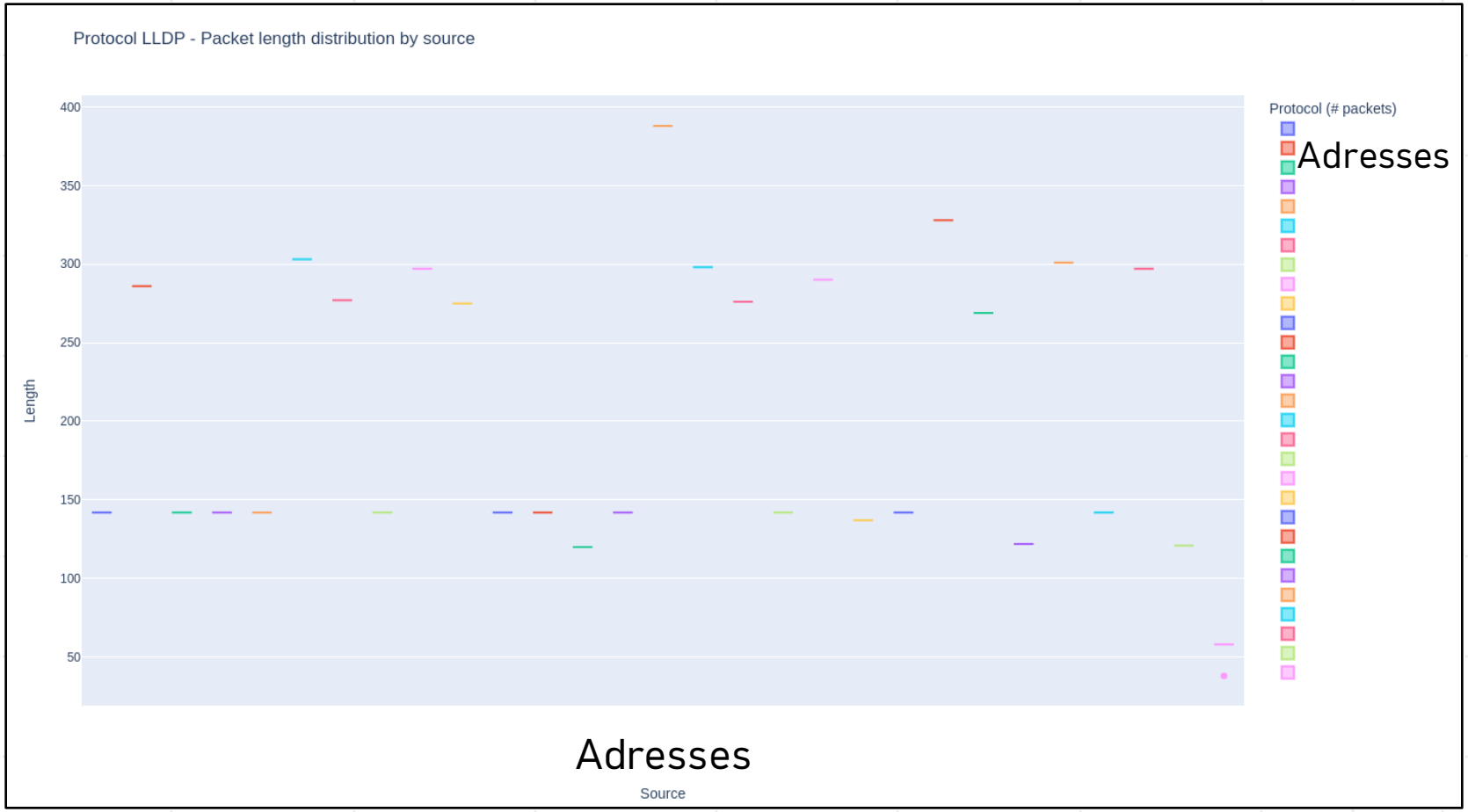
Vue d'ensemble → Check
Concept de base → Check
Information Exclusive → Check
Introduction d'autre Viz → Check



Visualisation Distribution

- Boxplot d'un protocoles, par addresse

Vue d'ensemble → Check
Concept de base → Check
Information Exclusive → Check
Introduction d'autre Viz → Check



Perspectives

Voyez-vous l'angle mort de cette méthodologie ?

Perspectives

Voyez-vous l'angle mort de cette méthodologie ?

- Rien pour examiner les labels de données labellisées !
- la raison : pas de labels dans le dataset de base.
- Prochain étape : éprouver méthodo sur données labellisées
  - peut-être super efficace également
  - peut-être pas
    - Adapter la méthodologie

Merci !