Détection collaborative d'intrusions dans le réseau

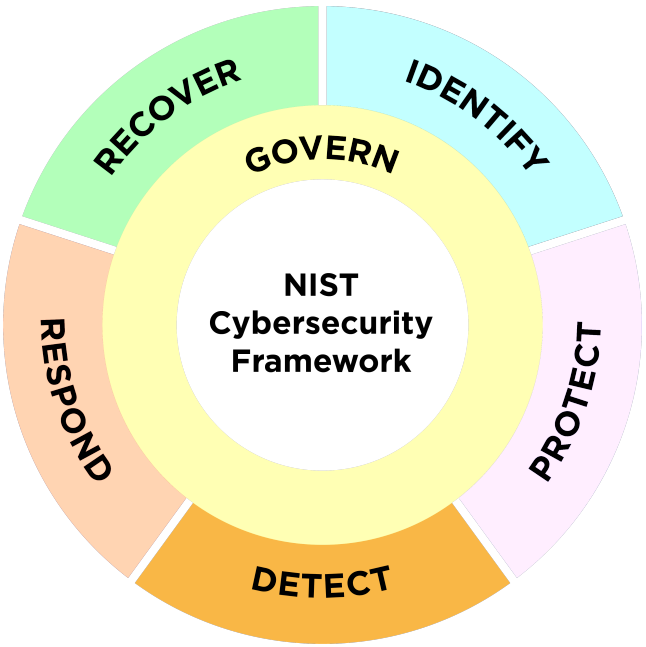
Gregory Blanc
SAMOVAR,Télécom SudParis
Institut Polytechnique de Paris

Cybersécurité dans l'ESR
Journée Thématique Nationale du CUME – 24 Mars 2026

$whoami

- maître de conférences à **Télécom SudParis**, une école de l'**IMT** et membre d'**IP Paris**
- responsable de la voie d'appronfondissement **SSR** *(Sécurité des Systèmes et Réseaux)*, formation labellisée *SecNumedu* et octroyant le titre **ESSI** délivré par l'**ANSSI**
- membre de l'équipe **SCN** *(Sécurité et Confiance Numérique)*, au laboratoire **SAMOVAR**
- membre associé du **LINCS**, laboratoire de R&I commun (Inria, IMT, Nokia, SU, SystemX)
- membre du comité de pilotage du colloque national **RESSI** *(Rendez-vous de la Recherche et de l'Enseignement en Sécurité des Systèmes d'Information)*
- membre du bureau du GT **SSLR** *(Sécurité des Systèmes, Logiciels et Réseaux)* du GDR CNRS Sécurité Informatique
- expert du comité technique du **PTCC** *(Programme de Transfert du Campus Cyber)*

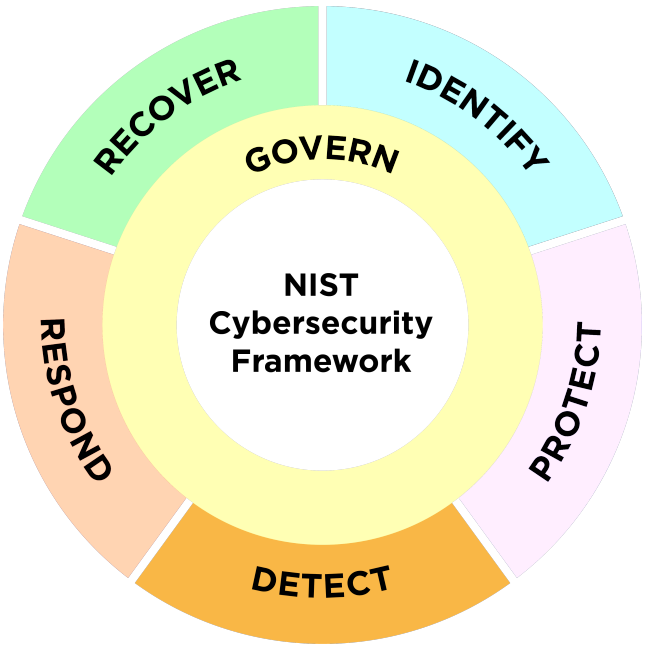
Fonction de détection (DE)



Detect au sens du référentiel CSF:

Les attaques et compromissions potentielles sont trouvées et analysées.

Fonction de détection (DE)

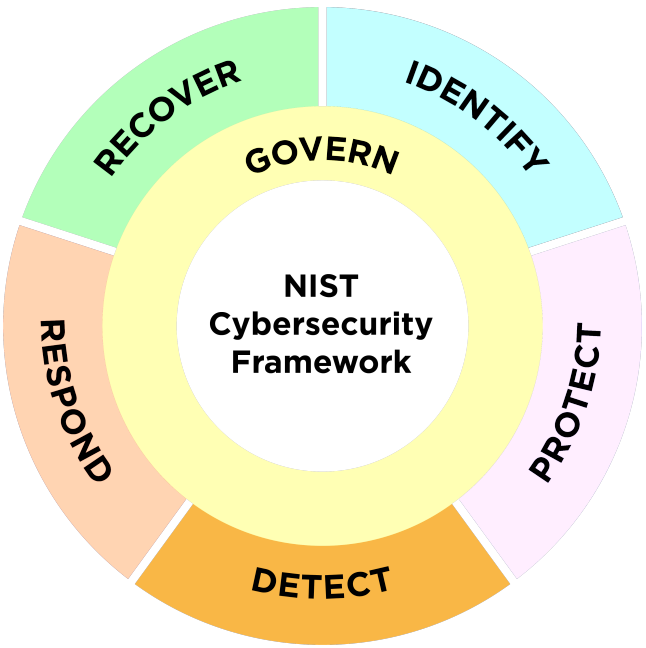


Detect au sens du référentiel CSF:

Les attaques et compromissions potentielles sont trouvées et analysées.

- Supervision continue
- Analyse des événements adverses

Fonction de détection (DE)



Detect au sens du référentiel CSF:

Les attaques et compromissions potentielles sont trouvées et analysées.

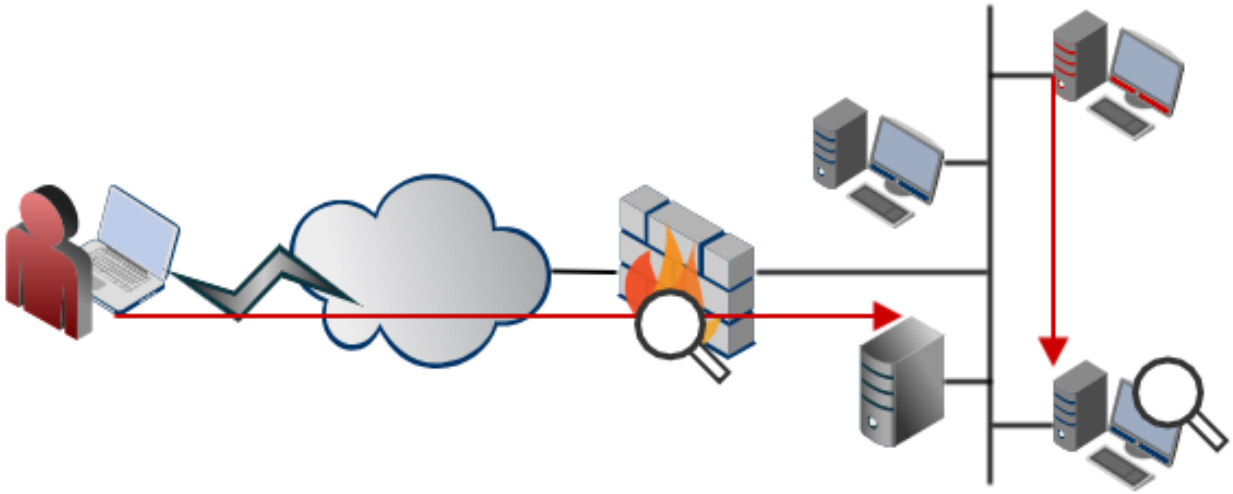
- **Supervision continue**

DE.CM-01 les réseaux et services sont supervisés

DE.CM-09 le matériek et le logiciel, les environnements d'exécution, et leurs données sont supervisés

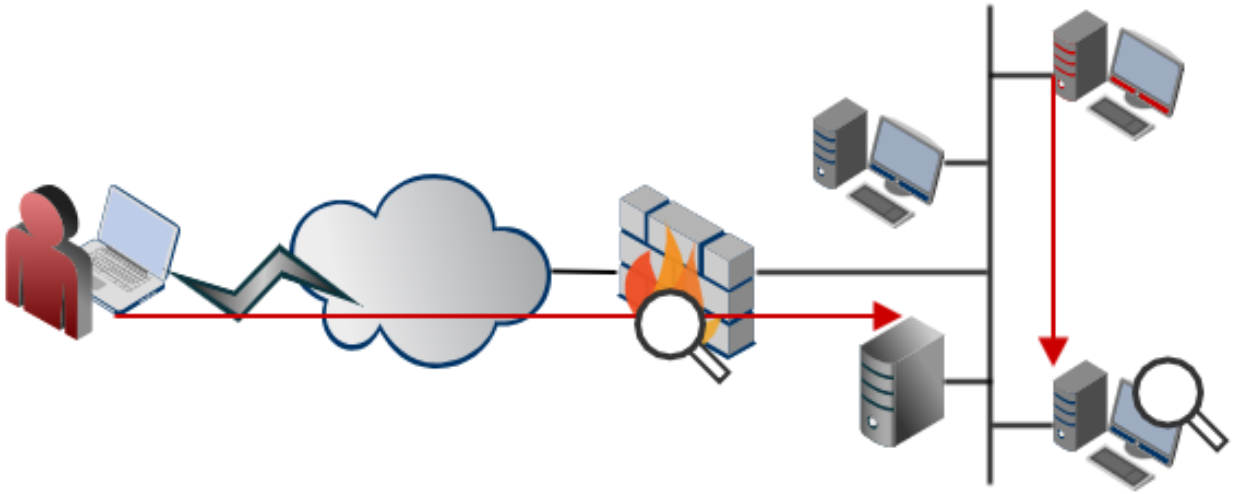
- Analyse des événements adverses

Détection d'intrusion



Lève une alerte lorsqu'une activité **suspecte** est identifiée

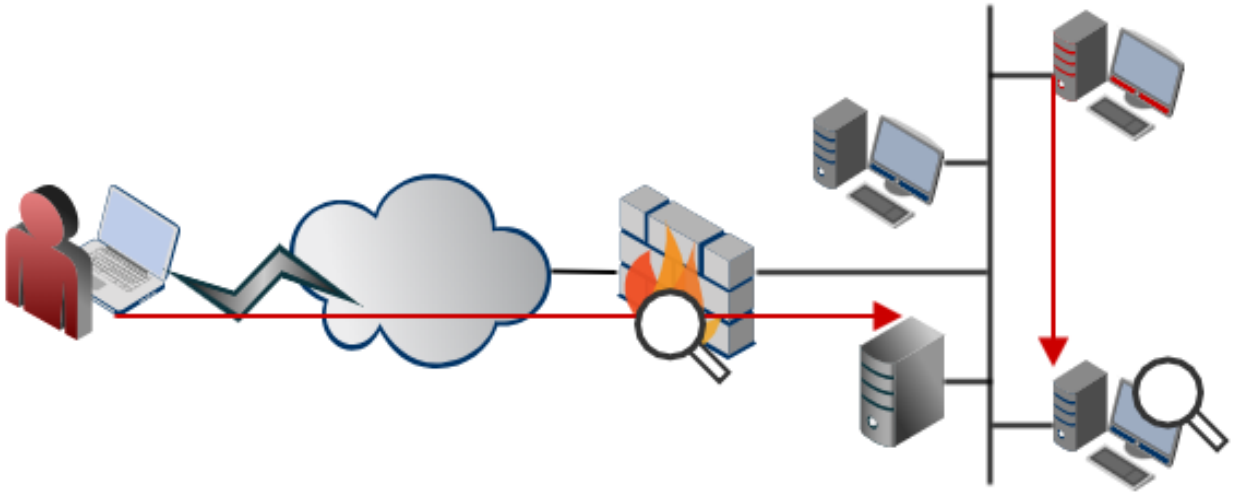
Détection d'intrusion



Lève une alerte lorsqu'une activité **suspecte** est identifiée

- Qu'est-ce qu'une activité suspecte ?

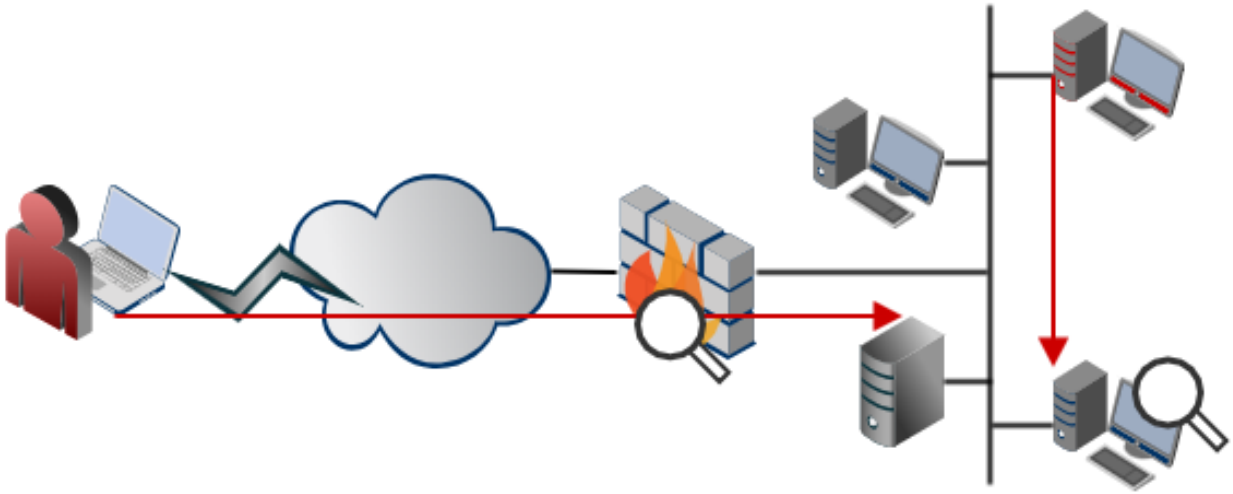
Détection d'intrusion



Lève une alerte lorsqu'une activité **suspecte** est identifiée

- Qu'est-ce qu'une activité suspecte ?
  - **un abus**: *activité connue pour être malveillante*
  - **une anomalie**: *une activité déviant de la normale*

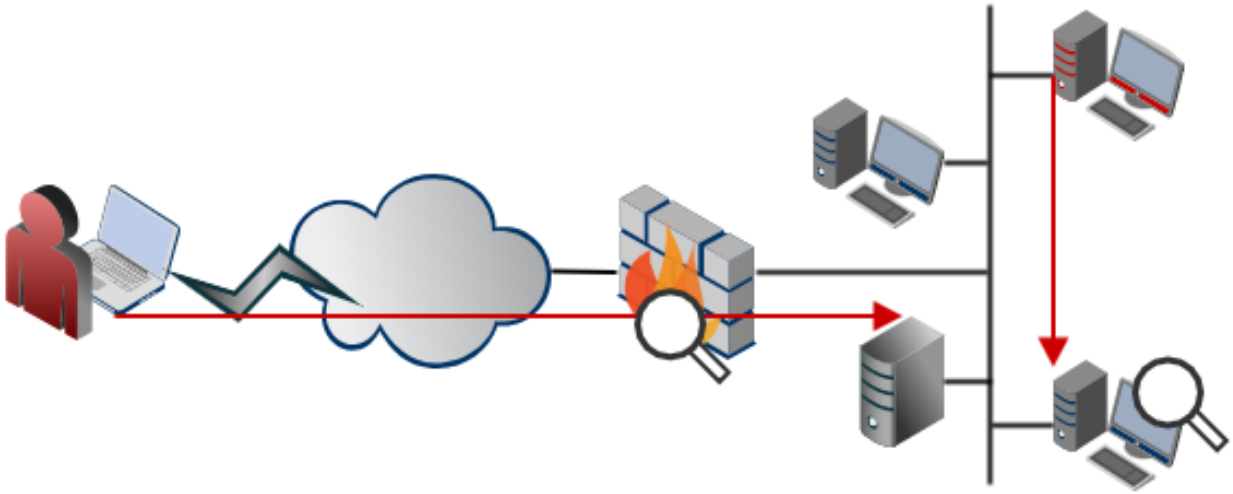
Détection d'intrusion



Lève une alerte lorsqu'une activité **suspecte** est identifiée

- Qu'est-ce qu'une activité suspecte ?
  - **un abus**: *activité connue pour être malveillante*
  - **une anomalie**: *une activité déviant de la normale*
- Comment capter les activités malveillantes ?

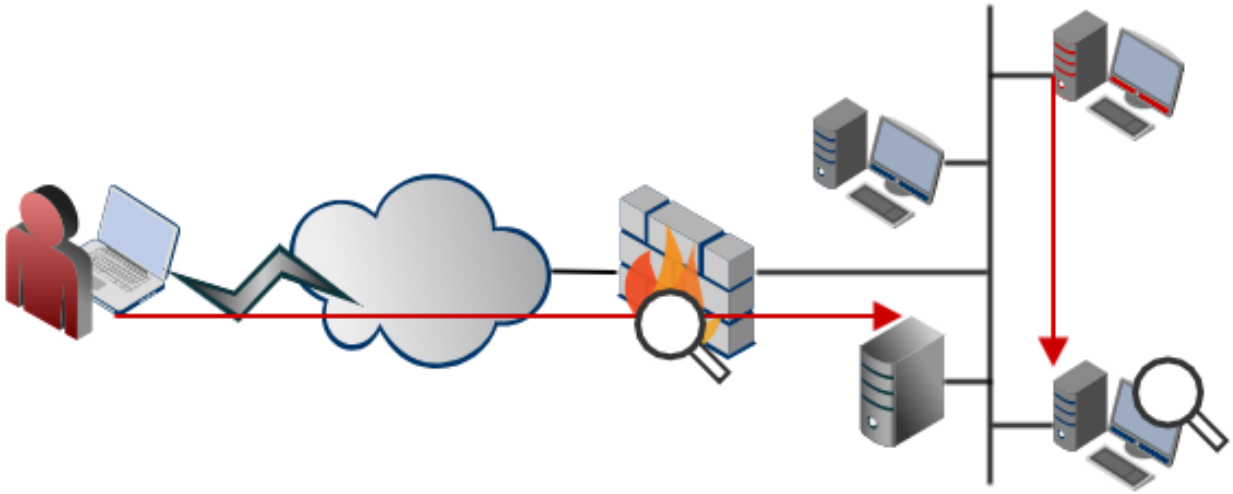
Détection d'intrusion



Lève une alerte lorsqu'une activité **suspecte** est identifiée

- Qu'est-ce qu'une activité suspecte ?
  - **un abus**: *activité connue pour être malveillante*
  - **une anomalie**: *une activité déviant de la normale*
- Comment capter les activités malveillantes ?
  - sur une machine : processus, logs, fichiers, etc.
  - dans le réseau : flux, en-tête de paquet, charge utile, etc.

Détection d'intrusion



Lève une alerte lorsqu'une activité **suspecte** est identifiée

- Qu'est-ce qu'une activité suspecte ?
  - **un abus**: *activité connue pour être malveillante*
  - **une anomalie**: *une activité déviant de la normale*
- Comment capter les activités malveillantes ?
  - sur une machine : processus, logs, fichiers, etc.
  - dans le réseau : flux, en-tête de paquet, charge utile, etc.

Plus les activités sont volumineuses, *plus* le traitement est long

Détection d'abus

Approche *principalement* des **signatures** d'attaques

Détection d'abus

Approche	*principalement* des **signatures** d'attaques
Éléments	en-têtes de paquets, statistiques de flux, connexions TCP, etc.

Détection d'abus

Approche	*principalement* des **signatures** d'attaques
Éléments	en-têtes de paquets, statistiques de flux, connexions TCP, etc.
Algorithmes	fouille de données et apprentissage statistique sur des jeux de données labellisées

Détection d'abus

Approche	*principalement* des **signatures** d'attaques
Éléments	en-têtes de paquets, statistiques de flux, connexions TCP, etc.
Algorithmes	fouille de données et apprentissage statistique sur des jeux de données labellisées
Défis	■ manque de jeux de données (existence, diversité, obsolescence, fiabilité) ■ fréquence de ré-entraînement

Détection d'abus

Approche	*principalement* des **signatures** d'attaques
Éléments	en-têtes de paquets, statistiques de flux, connexions TCP, etc.
Algorithmes	fouille de données et apprentissage statistique sur des jeux de données labellisées
Défis	■ manque de jeux de données (existence, diversité, obsolescence, fiabilité) ■ fréquence de ré-entraînement

Classification multi-classes

Détection d'abus

Approche	*principalement* des **signatures** d'attaques
Éléments	en-têtes de paquets, statistiques de flux, connexions TCP, etc.
Algorithmes	fouille de données et apprentissage statistique sur des jeux de données labellisées
Défis	■ manque de jeux de données (existence, diversité, obsolescence, fiabilité) ■ fréquence de ré-entraînement

Classification multi-classes

- Chaque classe encode un **motif**, similaire à une **signature**

Détection d'abus

Approche	*principalement* des **signatures** d'attaques
Éléments	en-têtes de paquets, statistiques de flux, connexions TCP, etc.
Algorithmes	fouille de données et apprentissage statistique sur des jeux de données labellisées
Défis	■ manque de jeux de données (existence, diversité, obsolescence, fiabilité) ■ fréquence de ré-entraînement

Classification multi-classes

- Chaque classe encode un **motif**, similaire à une **signature**
- L'apprentissage est **limité** aux classes d'attaque présentes dans le jeu d'entraînement

Détection d'abus

Approche	*principalement* des **signatures** d'attaques
Éléments	en-têtes de paquets, statistiques de flux, connexions TCP, etc.
Algorithmes	fouille de données et apprentissage statistique sur des jeux de données labellisées
Défis	■ manque de jeux de données (existence, diversité, obsolescence, fiabilité) ■ fréquence de ré-entraînement

Classification multi-classes

- Chaque classe encode un **motif**, similaire à une **signature**
- L'apprentissage est **limité** aux classes d'attaque présentes dans le jeu d'entraînement
- Soulage néanmoins **l'effort et le risque** de concevoir des signatures manuellement

Détection d'anomalies

Approche profils de comportements (*normaux*)

Détection d'anomalies

Approche profils de comportements (*normaux*)
Apprentissage supervisé, semi-supervisé et non-supervisé

Détection d'anomalies

Approche profils de comportements (*normaux*)

Apprentissage supervisé, semi-supervisé et non-supervisé

Défis

- *propreté* des jeux de données
- *exactitude* des comportements normaux
- *haut* taux de faux positifs

Détection d'anomalies

Approche profils de comportements (*normaux*)

Apprentissage supervisé, semi-supervisé et non-supervisé

Défis
- *propreté* des jeux de données
- *exactitude* des comportements normaux
- *haut* taux de faux positifs

Classification binaire

Détection d'anomalies

Approche profils de comportements (*normaux*)
Apprentissage supervisé, semi-supervisé et non-supervisé
Défis
- *propreté* des jeux de données
- *exactitude* des comportements normaux
- *haut* taux de faux positifs

Classification binaire

- Entraîner sur des données bénignes **uniquement** produit des motfis de comportement normal

Détection d'anomalies

Approche profils de comportements (*normaux*)
Apprentissage supervisé, semi-supervisé et non-supervisé
Défis
- *propreté* des jeux de données
- *exactitude* des comportements normaux
- *haut* taux de faux positifs

Classification binaire

- Entraîner sur des données bénignes **uniquement** produit des motfis de comportement normal
- Toute **déviation** du modèle entraîné déclenchent une détection

Détection d'anomalies

Approche profils de comportements (*normaux*)
Apprentissage supervisé, semi-supervisé et non-supervisé
Défis
- *propreté* des jeux de données
- *exactitude* des comportements normaux
- *haut* taux de faux positifs

Classification binaire

- Entraîner sur des données bénignes **uniquement** produit des motfis de comportement normal
- Toute **déviation** du modèle entraîné déclenchent une détection
- Manque de précision : une anomalie n'est pas **nécessairement** une malveillance

Détection d'anomalies

Approche profils de comportements (*normaux*)
Apprentissage supervisé, semi-supervisé et non-supervisé
Défis
- *propreté* des jeux de données
- *exactitude* des comportements normaux
- *haut* taux de faux positifs

Classification binaire

- Entraîner sur des données bénignes **uniquement** produit des motfis de comportement normal
- Toute **déviation** du modèle entraîné déclenchent une détection
- Manque de précision : une anomalie n'est pas **nécessairement** une malveillance

Mythe: *contrairement aux signatures, la détection d'anomalies utilise l'apprentissage machine (ML)*

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)
- 75% des opérateurs 5G remontent **1 à 6** failles par an (Nokia and GlobalData, 2022)

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)
- 75% des opérateurs 5G remontent **1 à 6** failles par an (Nokia and GlobalData, 2022)
- Dépenses en sécurité 5G **augmenteront** : de 4 Mds$ à 11 Mds$ (2029) (ABI Research, 2025)

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)
- 75% des opérateurs 5G remontent **1 à 6** failles par an (Nokia and GlobalData, 2022)
- Dépenses en sécurité 5G **augmenteront** : de 4 Mds$ à 11 Mds$ (2029) (ABI Research, 2025)

Mesures de sécurité de réseaux 5G

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)
- 75% des opérateurs 5G remontent **1 à 6** failles par an (Nokia and GlobalData, 2022)
- Dépenses en sécurité 5G **augmenteront** : de 4 Mds$ à 11 Mds$ (2029) (ABI Research, 2025)

Mesures de sécurité de réseaux 5G

- Authentification des appareils

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)
- 75% des opérateurs 5G remontent **1 à 6** failles par an (Nokia and GlobalData, 2022)
- Dépenses en sécurité 5G **augmenteront** : de 4 Mds$ à 11 Mds$ (2029) (ABI Research, 2025)

Mesures de sécurité de réseaux 5G

- Authentification des appareils
- Isolation des *slices* réseau

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)
- 75% des opérateurs 5G remontent **1 à 6** failles par an (Nokia and GlobalData, 2022)
- Dépenses en sécurité 5G **augmenteront** : de 4 Mds$ à 11 Mds$ (2029) (ABI Research, 2025)

Mesures de sécurité de réseaux 5G

- Authentification des appareils
- Isolation des *slices* réseau
- APIs sécurisées (entre slices et fonctions réseau)

Cas d'usage : opérateur de réseaux 5G

Besoins en sécurité

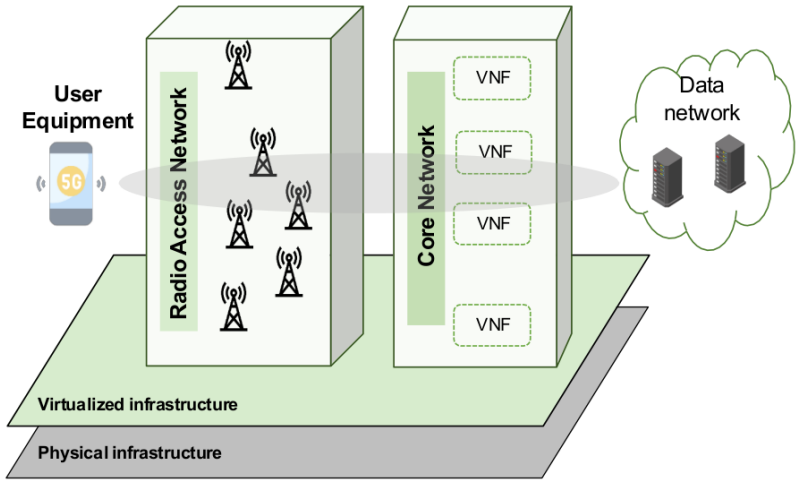
- L'architecture **virtualisée** et **cloudifiée** introduit de nouvelles attaques (ENISA, 2020)
- 75% des opérateurs 5G remontent **1 à 6** failles par an (Nokia and GlobalData, 2022)
- Dépenses en sécurité 5G **augmenteront** : de 4 Mds$ à 11 Mds$ (2029) (ABI Research, 2025)

Mesures de sécurité de réseaux 5G

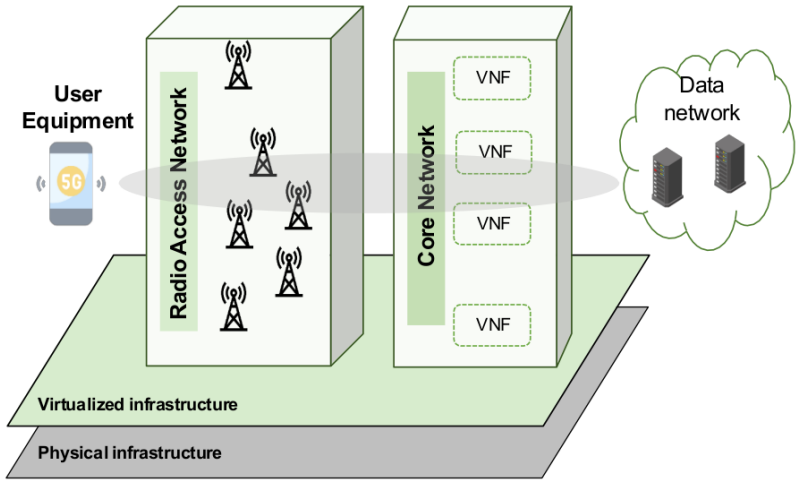
- Authentification des appareils
- Isolation des *slices* réseau
- APIs sécurisées (entre slices et fonctions réseau)

Les risques résiduels subsistent : *mauvaises configurations, appareils compromis, etc.*

Les défis de la détection en 5G [1]

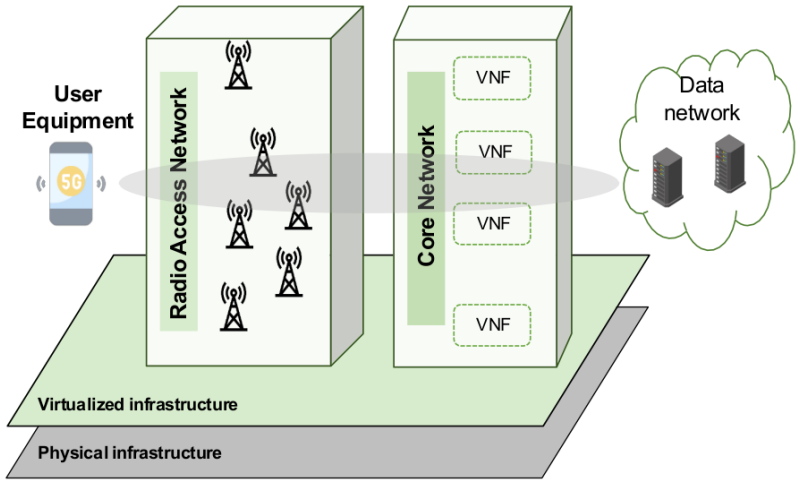


Les défis de la détection en 5G [1]



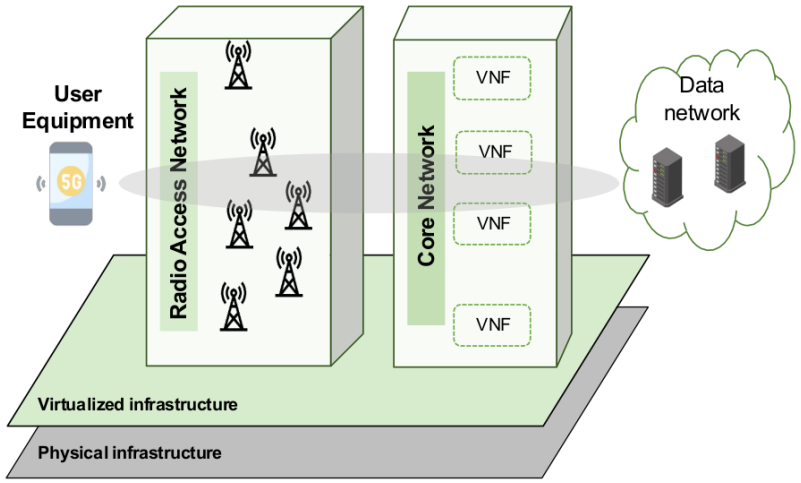
- Débit de données élevé

Les défis de la détection en 5G [1]



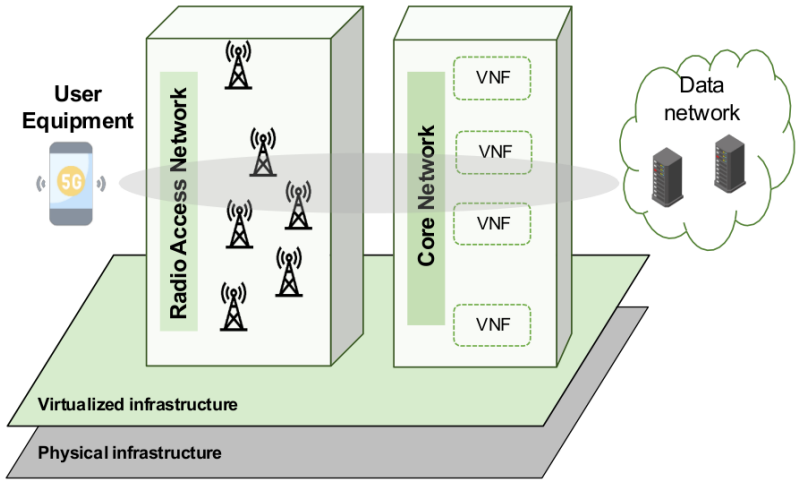
- Débit de données élevé → passage à l'échelle, (quasi) temps réel

Les défis de la détection en 5G [1]



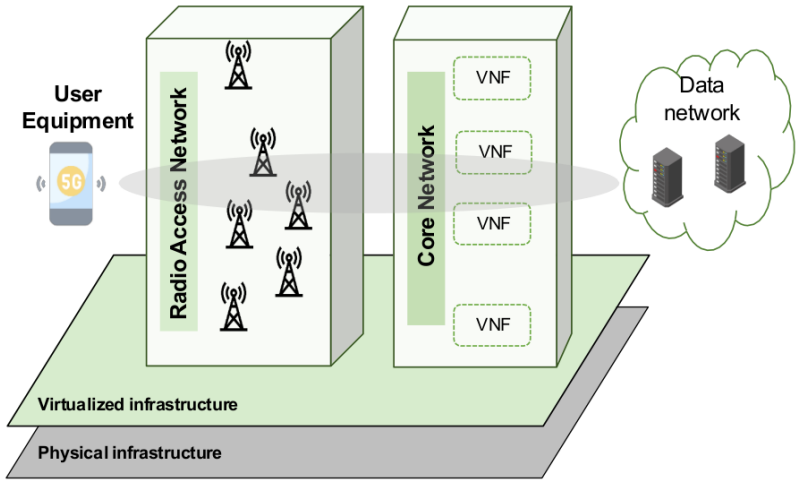
- Débit de données élevé → passage à l'échelle, (quasi) temps réel
- Fiabilité

Les défis de la détection en 5G [1]



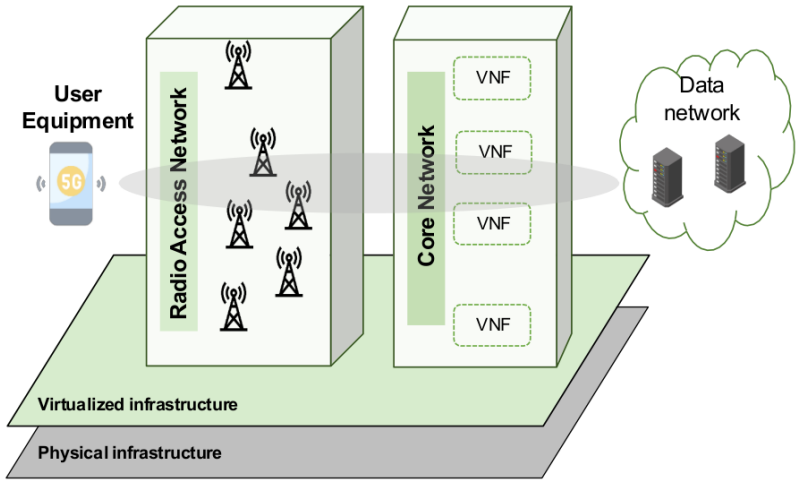
- Débit de données élevé $\rightarrow$ passage à l'échelle, (quasi) temps réel
- Fiabilité $\rightarrow$ interprétabilité, confiance

Les défis de la détection en 5G [1]



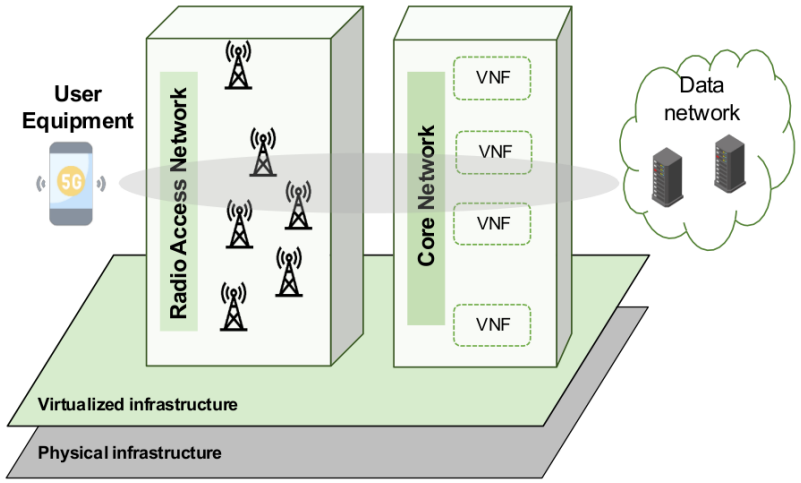
- Débit de données élevé → passage à l'échelle, (quasi) temps réel
- Fiabilité → interprétabilité, confiance
- Efficacité énergétique

Les défis de la détection en 5G [1]



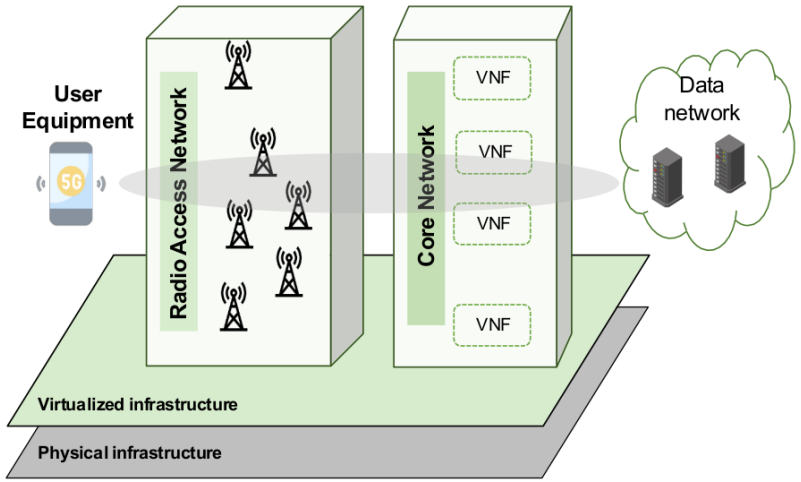
- Débit de données élevé → passage à l'échelle, (quasi) temps réel
- Fiabilité → interprétabilité, confiance
- Efficacité énergétique → localisation et complexité des modèles

Les défis de la détection en 5G [1]



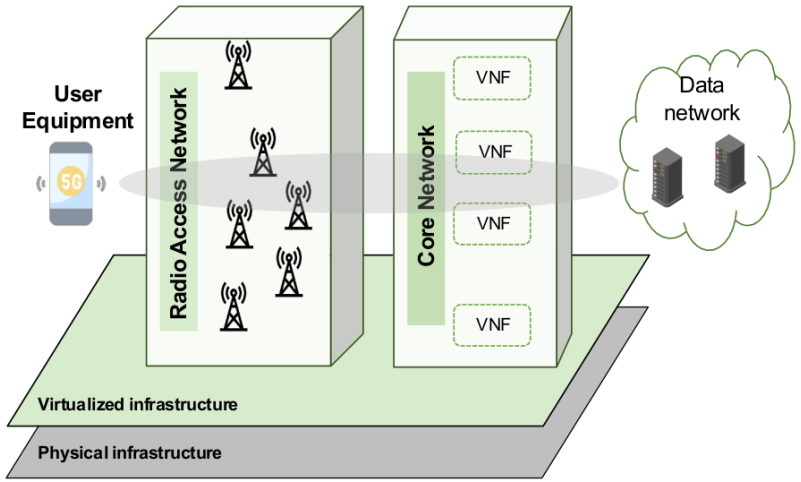
- Débit de données élevé → passage à l'échelle, (quasi) temps réel
- Fiabilité → interprétabilité, confiance
- Efficacité énergétique → localisation et complexité des modèles
- Hétérogénéité des données

Les défis de la détection en 5G [1]



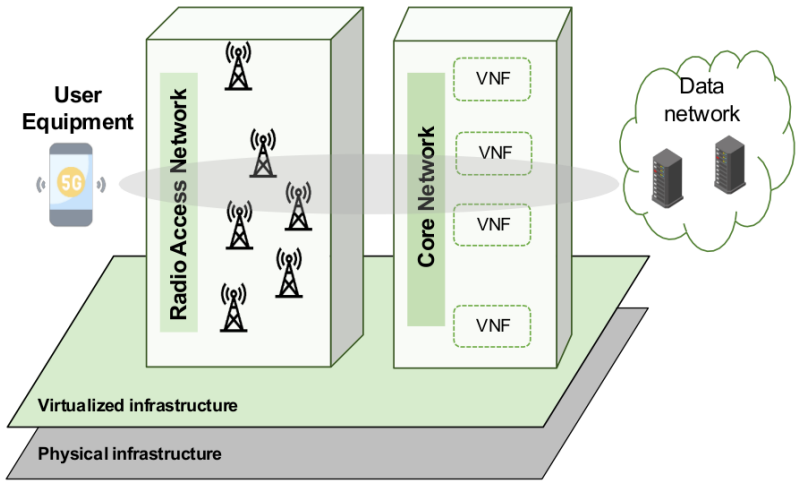
- Débit de données élevé → passage à l'échelle, (quasi) temps réel
- Fiabilité → interprétabilité, confiance
- Efficacité énergétique → localisation et complexité des modèles
- Hétérogénéité des données → adaptation à la diversité de comportements et distributions

Les défis de la détection en 5G [1]



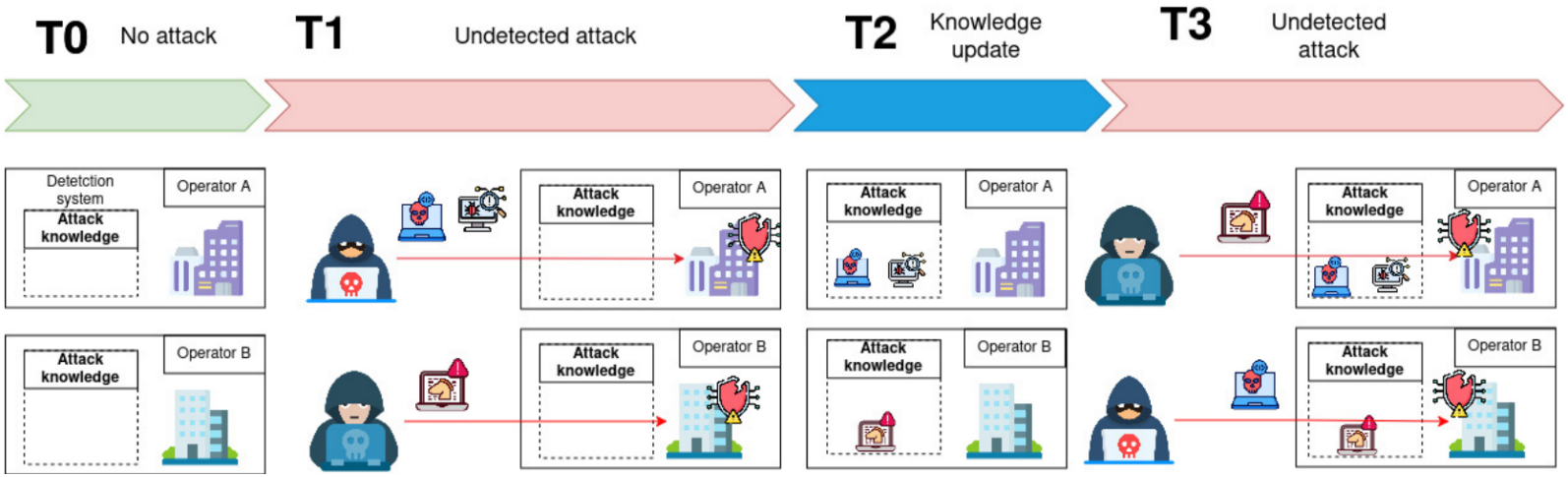
- Débit de données élevé → passage à l'échelle, (quasi) temps réel
- Fiabilité → interprétabilité, confiance
- Efficacité énergétique → localisation et complexité des modèles
- Hétérogénéité des données → adaptation à la diversité de comportements et distributions
- Mutualisation des ressources

Les défis de la détection en 5G [1]

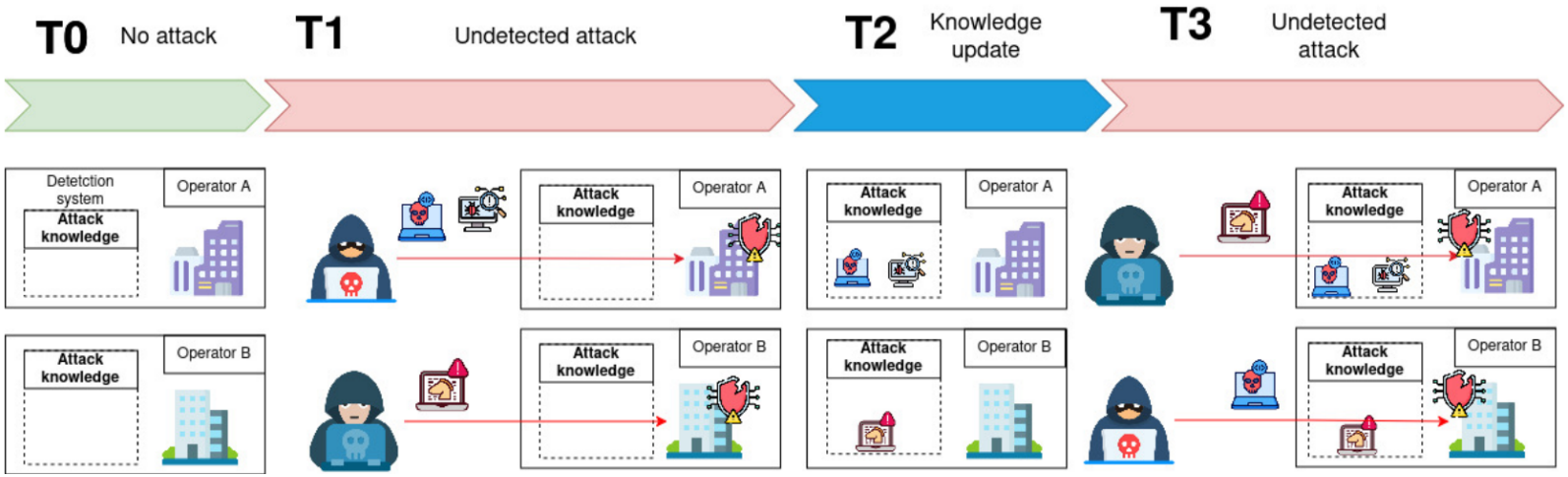


- Débit de données élevé → passage à l'échelle, (quasi) temps réel
- Fiabilité → interprétabilité, confiance
- Efficacité énergétique → localisation et complexité des modèles
- Hétérogénéité des données → adaptation à la diversité de comportements et distributions
- Mutualisation des ressources → collaboration vs. protection des données

Détection collaborative : motivation

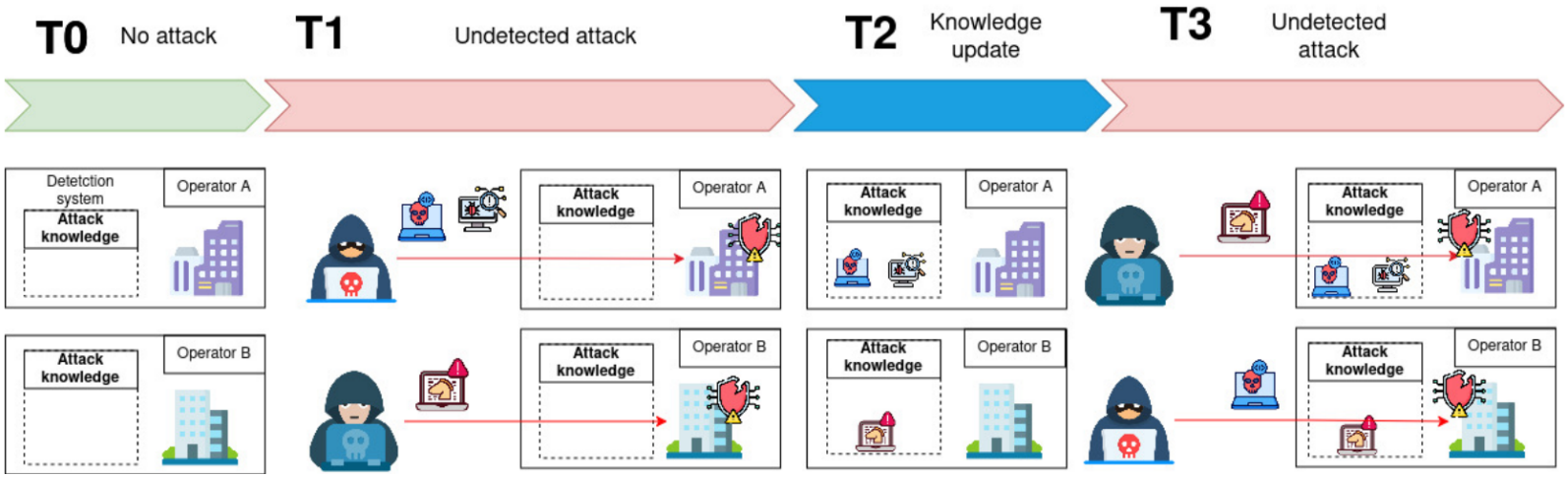


Détection collaborative : motivation



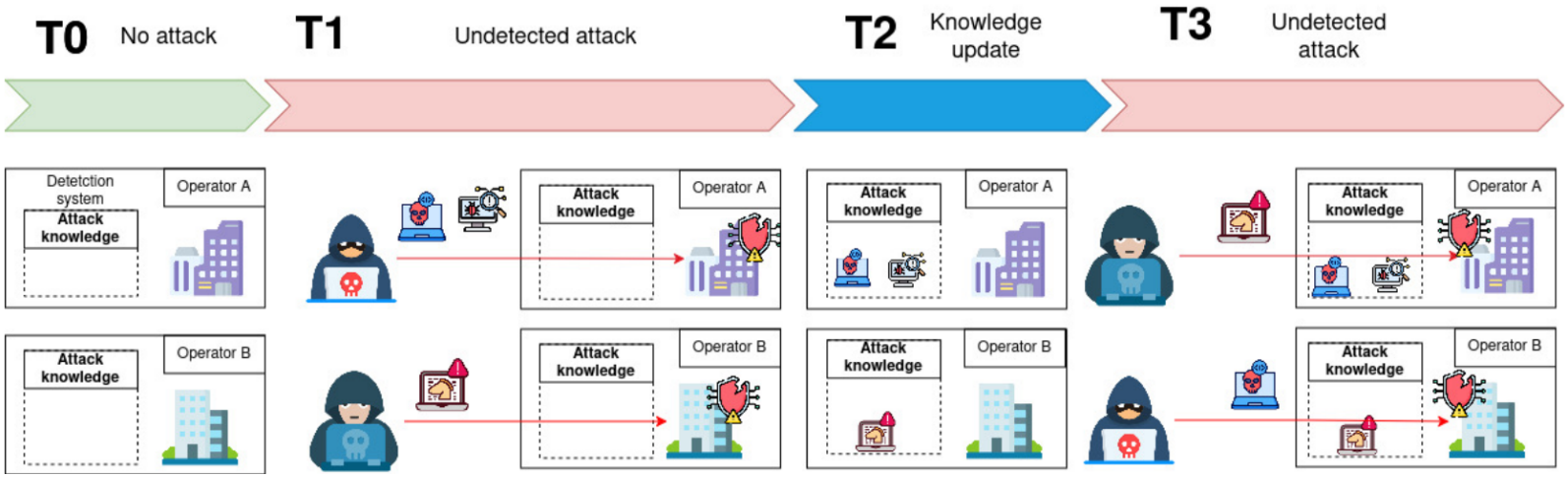
- Les attaques se propagent rapidement : réutilisation de vecteurs contre de nouvelles cibles

Détection collaborative : motivation



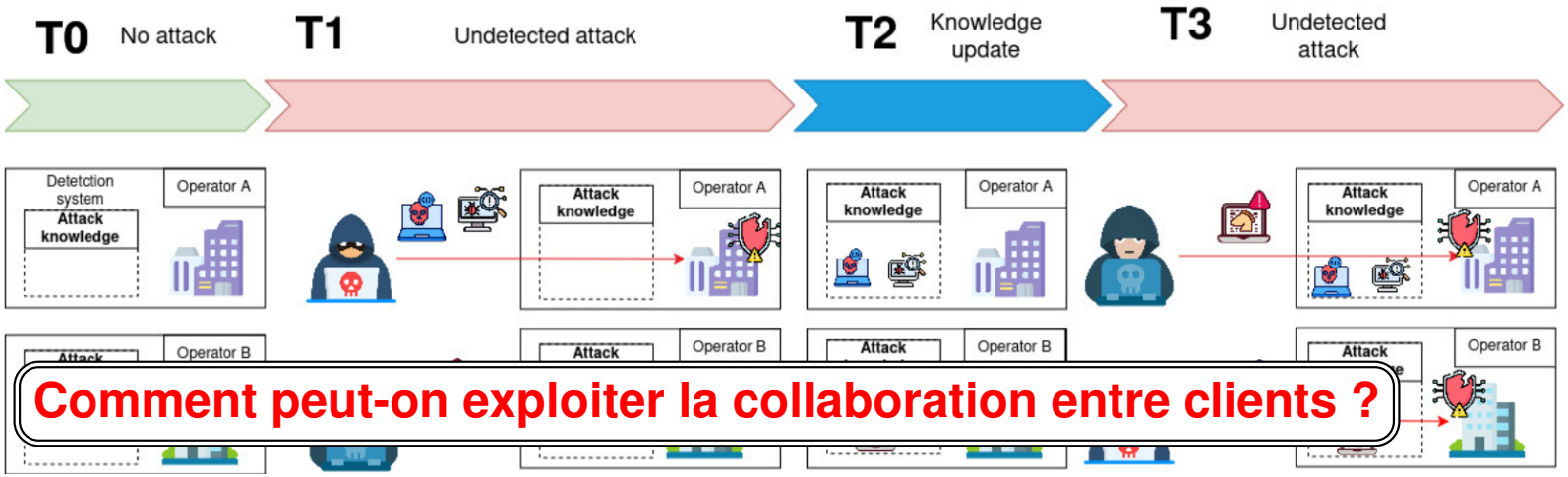
- Les attaques se propagent rapidement : réutilisation de vecteurs contre de nouvelles cibles
- Point d'observation central : les opérateurs peuvent superviser à travers les réseaux

Détection collaborative : motivation



- Les attaques se propagent rapidement : réutilisation de vecteurs contre de nouvelles cibles
- Point d'observation central : les opérateurs peuvent superviser à travers les réseaux
- Peu de temps pour réagir : dès l'attaque avérée, il faut partager des indicateurs

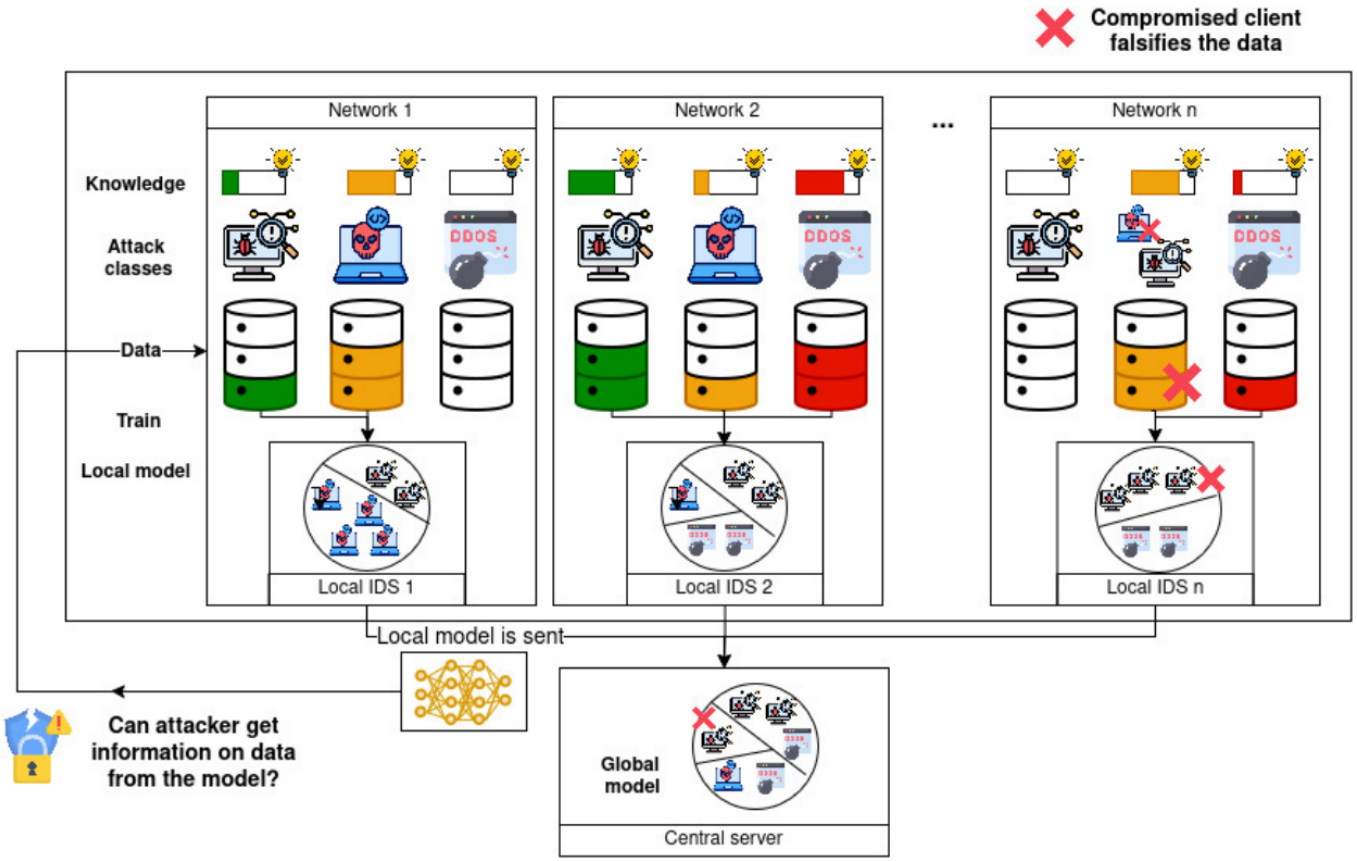
Détection collaborative : motivation



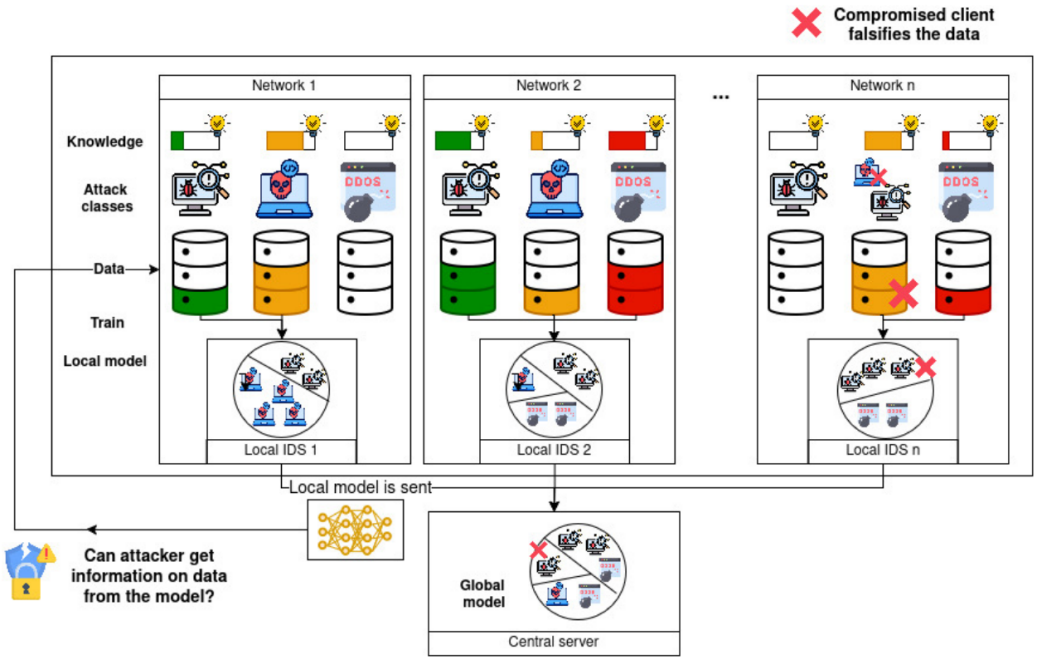
- Les attaques se propagent rapidement : réutilisation de vecteurs contre de nouvelles cibles
- Point d'observation central : les opérateurs peuvent superviser à travers les réseaux
- Peu de temps pour réagir : dès l'attaque avérée, il faut partager des indicateurs

En bref : l'apprentissage fédéré (FL)

IDS réseau collaboratif par apprentissage fédéré [2]



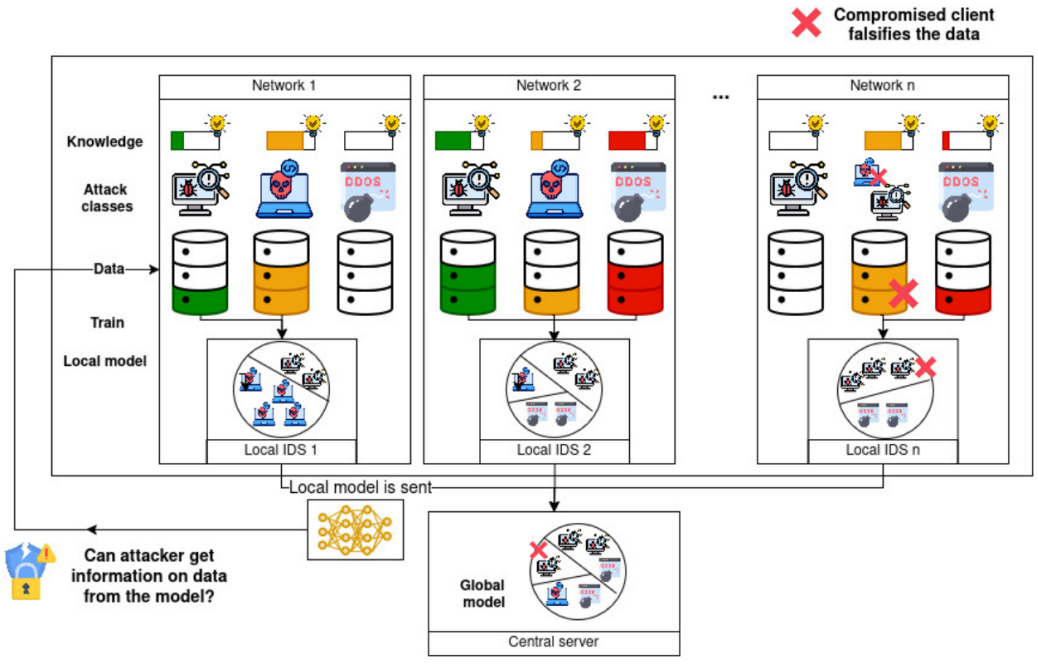
IDS réseau collaboratif par apprentissage fédéré [2]



Défis

- **Hétérogénéité :** les collaborateurs partagent des données très diverses

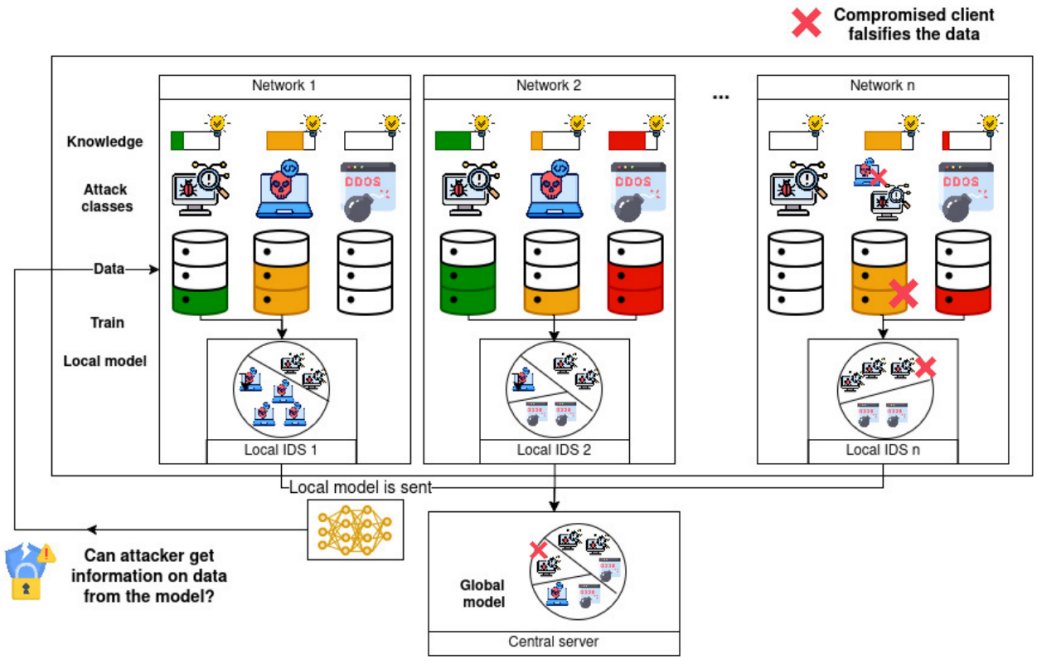
IDS réseau collaboratif par apprentissage fédéré [2]



Défis

- **Hétérogénéité :** les collaborateurs partagent des données très diverses
- **Protection des données sensibles :** le fournisseur de services est honnête-mais-curieux

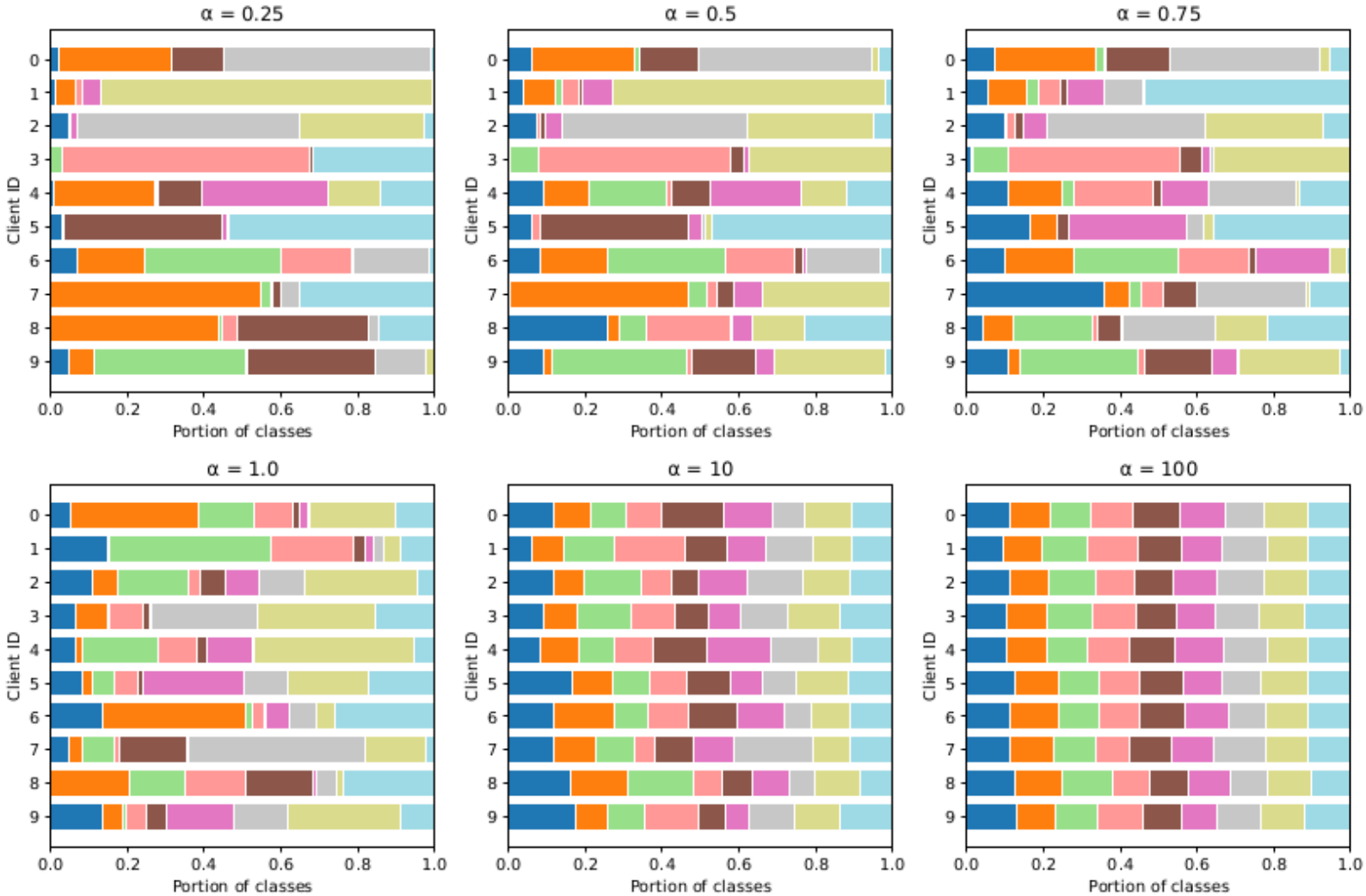
IDS réseau collaboratif par apprentissage fédéré [2]



Défis

- **Hétérogénéité :** les collaborateurs partagent des données très diverses
- **Protection des données sensibles :** le fournisseur de services est honnête-mais-curieux
- **Robustesse :** des clients compromis peuvent polluer les MAJ

Évaluation expérimentale : hétérogénéité

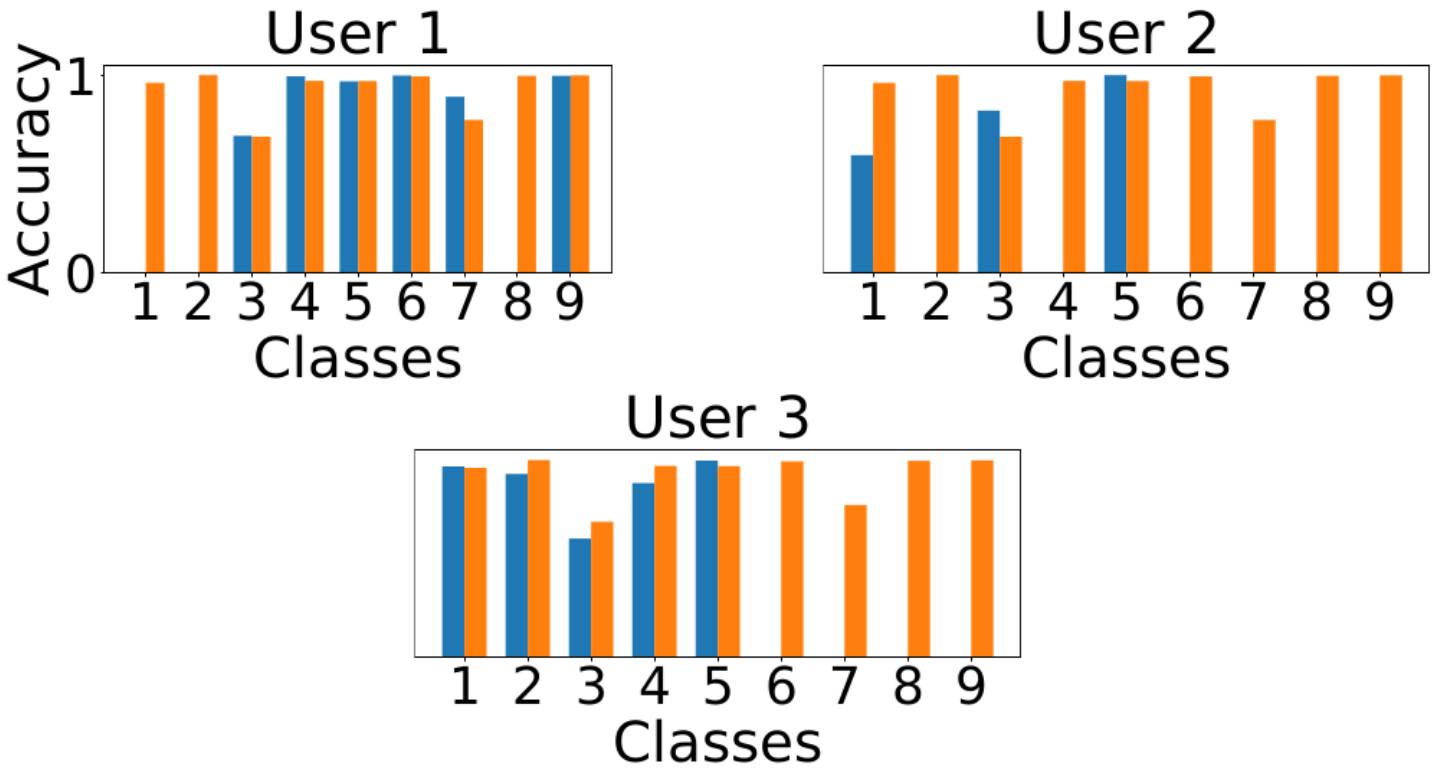


Évaluation expérimentale : résultats

Algorithm	X-IIoTID		
	α=0.75	α=0.50	α=0.25
Cerberus (using FedAvg)	52.23	56.52	53.57
FedProx-IDS	89.72	89.15	87.29
FPL-IDS	64.57	65.46	57.32
PROTEAN (Ours)	**92.67**	**93.62**	**93.43**

Note : $ACC = \frac{TP+TN}{P+N}$

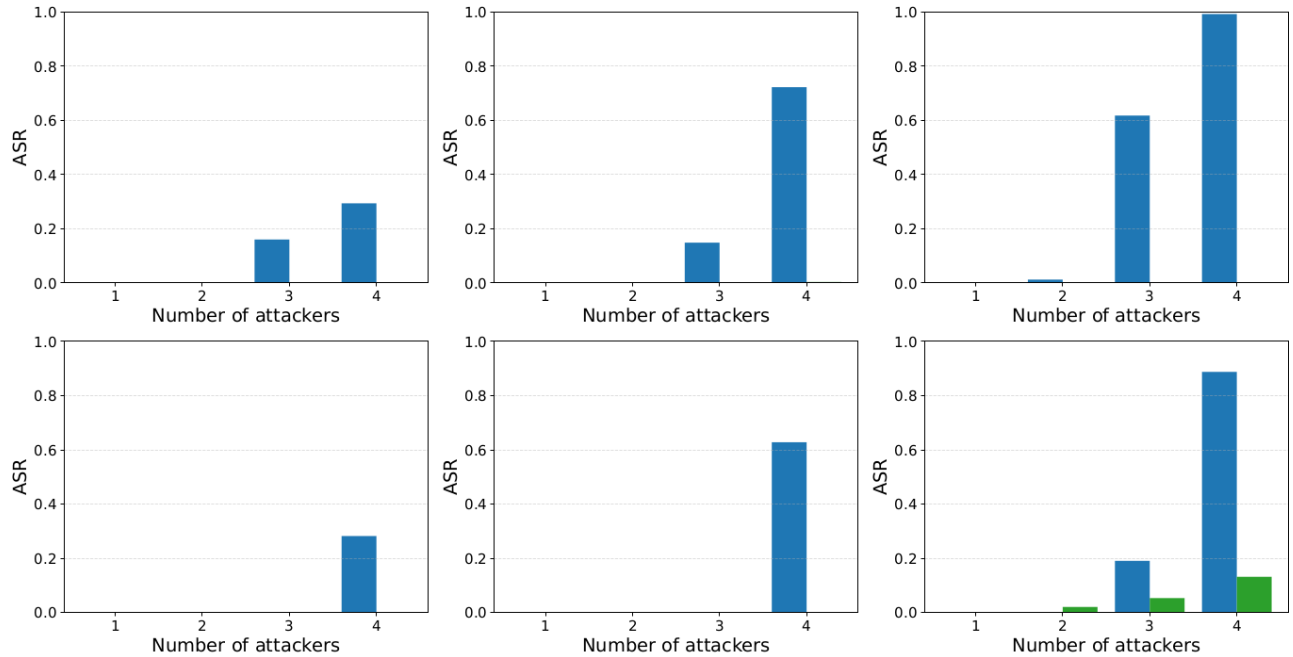
Évaluation expérimentale : partage de connaissances



Note : ACC par classe avec $\alpha = 0{,}25$

PROTEAN permet aux collaborateurs d'apprendre des motifs d'attaques inconnues partagées par d'autres clients.

Évaluation expérimentale : robustesse



(a) $\alpha = 0.75$ (b) $\alpha = 0.5$ (c) $\alpha = 0.25$

Note : Comparaison du taux de succès des attaques de basculement de labels (*label flipping*) face à Cerberus (FedAvg).

PROTEAN est nativement résistant aux attaques d'évasion par pollution des mises à jour.

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé
- les clients sont capables de labelliser leurs données

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé
- les clients sont capables de labelliser leurs données
- chaque attaque a été vu par au moins un des clients de la fédération

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé
- les clients sont capables de labelliser leurs données
- chaque attaque a été vu par au moins un des clients de la fédération

Les travaux futurs incluent :

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé
- les clients sont capables de labelliser leurs données
- chaque attaque a été vu par au moins un des clients de la fédération

Les travaux futurs incluent :

- la détection d'attaques *inconnues de tous*

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé
- les clients sont capables de labelliser leurs données
- chaque attaque a été vu par au moins un des clients de la fédération

Les travaux futurs incluent :

- la détection d'attaques *inconnues de tous*
- l'application de l'approche à la détection d'anomalies

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé
- les clients sont capables de labelliser leurs données
- chaque attaque a été vu par au moins un des clients de la fédération

Les travaux futurs incluent :

- la détection d'attaques *inconnues de tous*
- l'application de l'approche à la détection d'anomalies
- l'étude de la robustesse contre d'autres attaques

Limitations et perspectives

PROTEAN s'appuie sur quelques hypothèses :

- les clients ont détecté des attaques par le passé
- les clients sont capables de labelliser leurs données
- chaque attaque a été vu par au moins un des clients de la fédération

Les travaux futurs incluent :

- la détection d'attaques *inconnues de tous*
- l'application de l'approche à la détection d'anomalies
- l'étude de la robustesse contre d'autres attaques
- l'étude de contraintes opérationnelles (synchronisation, sélection de clients, équité, etc.)

Pour discuter

Contact

https://anr-grifin.telecom-sudparis.eu

gregory.blanc@telecom-sudparis.eu

Références

S. Chennoufi, G. Blanc, H. Jmila, and C. Kiennert, "Sok: Federated learning based network intrusion detection in 5g: Context, state of the art and challenges," in *Proceedings of the 19th International Conference on Availability, Reliability and Security, ARES 2024, Vienna, Austria, 30 July 2024 - 2 August 2024*, pp. 42:1–42:13, ACM, 2024.

S. Chennoufi, Y. Han, G. Blanc, E. De Cristofaro, and C. Kiennert, "PROTEAN: Federated Intrusion Detection in Non-IID Environments Through Prototype-Based Knowledge Sharing," in *Computer Security – ESORICS 2025* (V. Nicomette, A. Benzekri, N. Boulahia-Cuppens, and J. Vaidya, eds.), (Cham), pp. 103–125, Springer Nature Switzerland, 2026.

Remerciements

Ces travaux ont été financés par l'ANR via les projets GRIFIN (ANR-20-CE39-0011) et SuperviZ (ANR-22-PECY-0008) et BPIFrance via le projet Beyond5G.

Les figures sont issues de la thèse de Sara Chennoufi.