TP2

L'émergence de technologies avancées telles que le **Cloud Computing** et l'**Internet des objets (IoT)** permet aux utilisateurs de transmettre des données en temps réel. Cependant, cette capacité présente également des risques de sécurité, notamment des attaques susceptibles de compromettre le système.

Pour répondre à ce problème, nous souhaitons concevoir un **système multi-agents (SMA)** chargé de surveiller la présence de signaux caractéristiques d'attaques au cours de la transmission des données provenant des capteurs.

Pour atteindre cet objectif, il est nécessaire de réaliser les tâches suivantes :

1. **Configurer la bibliothèque JADE.**
2. **Créer les agents appropriés au sein du système à l'aide de la bibliothèque JADE**, à savoir :
    - Agent de prétraitement des données (*Data Pre-processing Agent*) ;
    - Agent de modélisation (*Model Agent*) ;
    - Agent de validation (*Validation Agent*).
3. **Entraîner des classificateurs** (arbres de décision, forêts aléatoires, SVM, réseaux de neurones, etc.) à partir de jeux de données étiquetés, tels que KDD-99 et NSL-KDD, afin de détecter les types d'attaques connus.
4. **Afficher les métriques d'évaluation** à chaque exécution ou évaluation du modèle.
5. **Afficher une vue d'ensemble de l'état global du système**, à l'aide de l'agent DF (*Directory Facilitator*) de JADE.

Remarques

- Le langage **Java** doit être utilisé pour développer les agents.
- Le langage **Python** doit être utilisé pour entraîner les modèles d'apprentissage automatique.
- La durée de réalisation de ce travail pratique est de **trois semaines**.

Jeux de données proposés

Jeu de données (année)	Environnement et trafic	Lien
KDD-99 / NSL-KDD (1999/2009)	Réseau local simulé (LAN), flux TCP/IP	NSL-KDD sur Kaggle
UNSW-NB15 (2015)	Environnement hybride combinant du trafic réel et synthétique, réseau de laboratoire	UNSW-NB15 sur Kaggle
CICIDS-2017 (2017)	Trafic réaliste d'un réseau d'entreprise, 25 utilisateurs	CICIDS-2017 sur Kaggle
BoT-IoT (2018)	Réseau IoT, banc d'essai de l'Université de Nouvelle-Galles du Sud à Canberra	BoT-IoT sur Kaggle
NF-UNSW-NB15-v3 (2020)	Représentation du jeu UNSW-NB15 sous forme de flux NetFlow	NF-UNSW-NB15 sur Kaggle
NF-UQ-NIDS / CIC-UQ-NIDS (2021-2022)	Jeux de données NetFlow fusionnés provenant de plusieurs sources	NF-UQ-NIDS sur Kaggle
OTIDS CANBus (2025)	Réseau CAN embarqué dans les véhicules	Jeu de données CAN Intrusion

Algorithmes d'apprentissage automatique

Les algorithmes suivants peuvent être utilisés pour entraîner et évaluer les modèles de détection des intrusions :

- Machines à vecteurs de support (**SVM — Support Vector Machine**) ;
- Arbres de décision (**DT — Decision Trees**) ;
- K plus proches voisins (**K-NN — K-Nearest Neighbors**) ;
- Forêts aléatoires (**Random Forest**) ;
- Régression logistique (*Logistic Regression*) ;
- Régression linéaire simple (*Simple Linear Regression*) ;
- Régression linéaire multiple (*Multiple Linear Regression*).

Ressource complémentaire : 100 Days of ML Code — SVM.