Détection d'intrusions fédérée et semi-supervisée pour l'IoT

O. Aouedi[1], K. Piamrat[1], G. Muller[2], Kamal Singh[2]

[1]LS2N, Universite de Nantes, BP 92208, 44322 Nantes Cedex 3, France
[2]Univ Jean Monnet, IOGS, CNRS, UMR 5516, LaHC, F - 42023 Saint-Etienne, France

guillaume.muller@univ-st-etienne.fr

Résumé

Le modèle `FLuIDS` combine apprentissages fédéré et semi-supervisé. De petits appareils (IoT) co-construisent un système de détection d'intrusions respectueux de la vie privée à moindre coût.

Mots-clés

Apprentissage Fédéré, Semi-Supervisé, IDS, IoT.

Abstract

`FLuIDS` combines federated and semi-supervised learning. Small devices (IoT) build an intrusion detection system that respects privacy and is cheaper.

Keywords

Federated Learning, Semi-Supervised, IDS, IoT.

Introduction

L'essor de l'Internet of Things (IoT) permet l'émergence de nombreuses applications : industrie du futur, « Smart Building »… Ces applications sont susceptibles d'être **attaquées** (déni de service…). Des systèmes de **détection d'intrusions** (IDS) doivent être mis en place. Le **Machine Learning** permet aux IDS d'offrir des solutions s'adaptant automatiquement à de nouveaux types d'attaques, sans intervention humaine, réduisant ainsi les **coûts** et le **temps** d'intervention. Dans un contexte IoT, l'application du Machine Learning « traditionnel » (très centralisé), demanderait des transferts de données et des calculs **trop lourds**. `FLuIDS` est un IDS adapté à l'IoT, par la combinaison : (i) du **Federated Learning** [2], pour *protéger la sensibilité des données* et pour *distribuer* les calculs et (ii) l'apprentissage **semi-supervisé**, pour exploiter au maximum les *données non-labellisées* (et éviter les coûts de labellisation).

Modèle `FLuIDS`

Dans `FLuIDS`, l'ensemble des appareils (nœuds terminaux + serveur) collaborent de manière fédérée pour construire un Auto-Encodeur (réseau de neurones non-supervisé) qui trouve la meilleure représentation latente pour les données. Le serveur extrait ensuite la partie « encodeur » (premières couches) et y adjoint une couche totalement connectée (FCN). Cette technique (« Split Learning ») résulte en un modèle entraînable de manière supervisée. Enfin, le serveur renvoie aux nœuds terminaux ce modèle, ce qui leur permet d'opérer eux-mêmes l'inférence (la détection d'intrusions). Ce processus est itéré continuellement, permettant un apprentissage « tout au long de la vie ».

Résultats

Pour vérifier la pertinence et l'efficacité de `FLuIDS`, nous l'avons testé sur deux jeux de données de référence dans le domaine : UNSW-NB15 et SCADA Gas Pipeline.

Coût de communication : `FLuIDS` permet de réduire les échanges de données tout en limitant la perte de performances par rapport à un modèle « traditionnel » centralisé (baisse de 20 % à 99 % suivant les configurations).

Performances face à des modèles supervisés : `FLuIDS` montre des performances légèrement supérieures (5 % à 10 %) classer le trafic normal et globalement équivalentes (0 % à 5 % meilleures) pour les attaques à celles de modèles entièrement supervisés (MLP, Forêt Aléatoires, Arbres de Décision, SVM). Ces derniers semblent plus efficaces pour extraire l'information des données labellisées, mais sont incapables d'exploiter les données non-labellisées.

Performances avec/sans fédération : `FLuIDS` se montre légèrement plus performant (≈3 %) qu'une version non-fédérée de lui-même (i.e. un modèle semi-supervisé mais centralisé), ce qui tend à montrer qu'une répartition aléatoire des données (potentiellement non I.I.D.) sur les nœuds force le modèle à mieux généraliser que quand il a directement accès à toutes les données.

Perspectives

`FLuIDS` démontre la viabilité de l'apprentissage fédérée pour l'IDS dans l'IoT. Un modèle non-supervisé plus adapté à la détection d'anomalies (Auto-Encodeur Variationnel) et/ou une fonction d'agrégation plus évoluée (e.g. SCAFFOLD [1]) permettrait certainement d'obtenir des taux de classification encore meilleurs. Enfin, l'accélération du processus d'entraînement des réseaux de neurones (par des alternatives à la rétro-propagation, ou des puces spécialisées – FPGA) permettrait l'exécution de `FLuIDS` sur de vraiment très petits appareils.

Références

[1] S.P. KARIMIREDDY et al. "SCAFFOLD : Stochastic controlled averaging for federated learning". In : *International Conference on Machine Learning*. PMLR. 2020, p. 5132-5143.

[2] B. MCMAHAN et al. "Communication-efficient learning of deep networks from decentralized data". In : *Artificial intelligence and statistics*. PMLR. 2017, p. 1273-1282.