COLLEGES BRETAGNE LOIRE | SCIENCES POUR L'INGENIEUR ET LE NUMERIQUE

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

THÈSE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPÉRIEURE
MINES-TÉLÉCOM ATLANTIQUE BRETAGNE
PAYS DE LA LOIRE – IMT ATLANTIQUE

ÉCOLE DOCTORALE N° 648
Sciences pour l'Ingénieur et le Numérique
Spécialité : *Mathématiques et Sciences et Technologies de l'Information et de la Communication*

Par
Léo LAVAUR

L'Apprentissage Fédéré comme Outil pour la Détection Collaborative d'Intrusions

Thèse présentée et soutenue à Rennes, le XX septembre 2024
Unité de recherche : IRISA (UMR 6074), SOTERN

Rapporteurs avant soutenance :

Anne-Marie Kermarrec	Professeure à l'Université Polytechnique Fédérales de Lausanne (EPFL)
Éric Totel	Professeur à Télécom SudParis

Composition du Jury :

Attention, en cas d'absence d'un des membres du Jury le jour de la soutenance, la composition du jury doit être revue pour s'assurer quelle est conforme et devra être répercutée sur la couverture de thèse

Président :	*À compléter après la soutenance.*	
Examinateurs :	Sonia Ben Mokhtar	Directrice de Recherche CNRS au laboratoire LIRIS
	Pierre-François Gimenez	Maître de Conférence à CentraleSupelec
	Vincent Nicomette	Professeur à l'INSA de Toulouse
	Fabien AUTREL	Ingénieur de Recherche à IMT Atlantique
	Marc-Oliver PAHL	Directeur d'Étude à IMT Atlantique
Dir. de thèse :	Yann BUSNEL	Directeur de la Recherche et de l'Innovation (DRI) à IMT Nord Europe

Invité(s) :

Prénom NOM	Fonction et établissement d'exercice

Abstracts

Résumé

La collaboration entre les différents acteurs de la cybersécurité est essentielle pour lutter contre des attaques de plus en plus sophistiquées et nombreuses. Pourtant, les organisations sont souvent réticentes à partager leurs données, par peur de compromettre leur confidentialité, et ce même si cela pourrait d'améliorer leurs modèles de détection d'intrusions. L'apprentissage fédéré est un paradigme récent en apprentissage automatique qui permet à des clients distribués d'entraîner un modèle commun sans partager leurs données. Ces propriétés de collaboration et de confidentialité en font un candidat idéal pour des applications sensibles comme la détection d'intrusions. Si un certain nombre d'applications ont montré qu'il est, en effet, possible d'entraîner un modèle unique sur des données distribuées de détection d'intrusions, peu se sont intéressées à l'aspect collaboratif de ce paradigme. En plus de l'aspect collaboratif, d'autres problématiques apparaissent dans ce contexte, telles que l'hétérogénéité des données des différents participants ou la gestion de participants non fiables. Dans ce manuscrit, nous explorons l'utilisation de l'apprentissage fédéré pour construire des systèmes collaboratifs de détection d'intrusions. En particulier, nous explorons l'impact de la qualité des données dans des contextes hétérogènes, certains types d'attaques par empoisonnement, et proposons des outils et des méthodologies pour améliorer l'évaluation de ce type d'algorithmes distribués.

Abstract

Collaboration between different cybersecurity actors is essential to fight against increasingly sophisticated and numerous attacks. However, stakeholders are often reluctant to share their data, fearing confidentiality and privacy issues, although it would improve their intrusion detection models. Federated learning is a recent paradigm in machine learning that allows distributed clients to train a common model without sharing their data. These properties of collaboration and confidentiality make it an ideal candidate for sensitive applications such as intrusion detection. While several applications have shown that it is indeed possible to train a single model on distributed intrusion detection data, few have focused on the collaborative aspect of this paradigm. In addition to the collaborative aspect, other challenges arise in this context, such as the heterogeneity of the data between different participants or the management of untrusted contributions. In this manuscript, we explore the use of federated learning to build collaborative intrusion detection systems. In particular, we explore the impact of data quality in heterogeneous contexts, some types of poisoning attacks, and propose tools and methodologies to improve the evaluation of these types of distributed algorithms.

ACKNOWLEDGEMENTS

Table of Contents

Part I

Federated Learning to build CIDSs

Chapter 2

Preliminaries

Contents

This chapter provides the necessary background on Machine Learning (ML) for intrusion detection, the inner of Federated Learning (FL), and the related threats.

2.1 Machine Learning for Intrusion Detection

2.2 Fundamentals of Federated Learning

2.3 Threats against Federated Learning

Part II

Quantifying the Limitations of FIDSs

Part III

Providing Solutions

List of Figures

List of Tables

COLLEGES SCIENCES
BRETAGNE POUR L'INGENIEUR
LOIRE ET LE NUMERIQUE

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Titre : L'Apprentissage Fédéré comme Outil pour la Détection Collaborative d'Intrusions

Mot clés : apprentissage automatique, apprentissage fédéré, détection d'intrusions, collaboration, confiance

Résumé : La collaboration entre les différents acteurs de la cybersécurité est essentielle pour lutter contre des attaques de plus en plus sophistiquées et nombreuses. Pourtant, les organisations sont souvent réticentes à partager leurs données, par peur de compromettre leur confidentialité, et ce même si cela pourrait d'améliorer leurs modèles de détection d'intrusions. L'apprentissage fédéré est un paradigme récent en apprentissage automatique qui permet à des clients distribués d'entraîner un modèle commun sans partager leurs données. Ces propriétés de collaboration et de confidentialité en font un candidat idéal pour des applications sensibles comme la détection d'intrusions. Si un certain nombre d'applications ont montré qu'il est, en effet, possible d'entraîner un modèle unique sur des données distribuées de détection d'intrusions, peu se sont intéressées à l'aspect collaboratif de ce paradigme. En plus de l'aspect collaboratif, d'autres problématiques apparaissent dans ce contexte, telles que l'hétérogénéité des données des différents participants ou la gestion de participants non fiables. Dans ce manuscrit, nous explorons l'utilisation de l'apprentissage fédéré pour construire des systèmes collaboratifs de détection d'intrusions. En particulier, nous explorons l'impact de la qualité des données dans des contextes hétérogènes, certains types d'attaques par empoisonnement, et proposons des outils et des méthodologies pour améliorer l'évaluation de ce type d'algorithmes distribués.

Title: On Federated Learning as a Framework for Collaborative Intrusion Detection

Keywords: machine learning, federated learning, intrusion detection, collaboration, trust

Abstract: Collaboration between different cybersecurity actors is essential to fight against increasingly sophisticated and numerous attacks. However, stakeholders are often reluctant to share their data, fearing confidentiality and privacy issues, although it would improve their intrusion detection models. Federated learning is a recent paradigm in machine learning that allows distributed clients to train a common model without sharing their data. These properties of collaboration and confidentiality make it an ideal candidate for sensitive applications such as intrusion detection. While several applications have shown that it is indeed possible to train a single model on distributed intrusion detection data, few have focused on the collaborative aspect of this paradigm. In addition to the collaborative aspect, other challenges arise in this context, such as the heterogeneity of the data between different participants or the management of untrusted contributions. In this manuscript, we explore the use of federated learning to build collaborative intrusion detection systems. In particular, we explore the impact of data quality in heterogeneous contexts, some types of poisoning attacks, and propose tools and methodologies to improve the evaluation of these types of distributed algorithms.