RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

Université Ferhat Abbas Sétif 1

Faculté des Sciences

Département d'informatique

MEMOIRE DE MASTER

Outil d'apprentissage automatique WEKA pour le système de détection d'intrusion

Réalisé par :

Yessad Djaafer

Bouzidi Aymene

Soutenu le 16/06/2025 devant le jury :

Dr. Frihia Hamza	UFAS 1	Encadrant
Dr. ZERGUINE Nadia	UFAS 1	Présidente
Dr. NASRI Khaled	UFAS 1	Examinateur

Année universitaire : 2024/2025

Résumé

Ces dernières années, la prolifération des cybermenaces a fait des systèmes de détection d'intrusion un élément essentiel de la cybersécurité. Les approches IDS traditionnelles, notamment celles basées sur des signatures prédéfinies, se heurtent à des limites pour détecter les attaques émergentes et sophistiquées. Ce mémoire explore l'intégration de techniques d'apprentissage automatique à l'outil WEKA, en se concentrant sur l'amélioration des capacités de détection grâce à des classificateurs intelligents.

Ce mémoire comprend une analyse détaillée de divers algorithmes d'apprentissage supervisé, appliqués au jeu de données NSL-KDD, une référence dans le domaine de la détection d'intrusion. Une contribution significative de ces travaux est le développement d'une solution hybride permettant l'utilisation de réseaux de neurones convolutifs 1D-CNN au sein de l'environnement WEKA. Cette intégration a été réalisée grâce à la conception d'une interface personnalisée reliant WEKA à des bibliothèques d'apprentissage profond basées sur Python, surmontant ainsi les limitations inhérentes à WEKA pour le traitement de données.

Le cadre proposé facilite l'apprentissage et l'évaluation fluides des modèles en combinant l'interface conviviale de WEKA avec la puissance de l'apprentissage profond. Cette approche ouvre la voie à la création de systèmes de détection d'intrusion plus robustes et adaptatifs. Les travaux futurs pourraient inclure l'extension de la solution à des scénarios de détection d'intrusion en temps réel et son test sur des ensembles de données plus diversifiés.

Mots clés : systèmes de détection d'intrusion, WEKA, apprentissage automatique, apprentissage profond, NSL-KDD, CNN.

Abstract

In recent years, the proliferation of cyber threats has made intrusion detection systems (IDS) a vital component of cybersecurity. Traditional IDS approaches, particularly those based on predefined signatures, face limitations when it comes to detecting emerging and sophisticated attacks. This research explores the integration of machine learning techniques into the WEKA tool, focusing on enhancing detection capabilities through intelligent classifiers.

The study involves a detailed analysis of various supervised learning algorithms, applied to the NSL-KDD dataset, a benchmark in the field of intrusion detection. A significant contribution of this work is the development of a hybrid solution enabling the use of 1D Convolutional Neural Networks 1D-CNN within the WEKA environment. This integration was achieved by designing a custom interface that connects WEKA with Python-based deep learning libraries, overcoming WEKA's inherent limitations for processing data.

The proposed framework facilitates seamless model training and evaluation by combining WEKA's user-friendly interface with the power of deep learning. This approach paves the way for building more robust and adaptive intrusion detection systems. Future work may include extending the solution to real-time intrusion detection scenarios and testing it on more diverse datasets.

Keywords: intrusion detection systems, WEKA, machine learning, Deep learning, NSL-KDD, CNN.

ملخص

في السنوات الأخيرة، أدى انتشار التهديدات السيبرانية إلى جعل أنظمة كشف التسلل (IDS) مكونًا حيويًا للأمن السيبراني. تواجه مناهج IDS التقليدية، وخاصة تلك القائمة على التوقيعات المحددة مسبقًا، قيودًا فيما يتعلق باكتشاف الهجمات الناشئة والمعقدة. يستكشف هذا البحث دمج تقنيات التعلم الآلي في أداة WEKA، مع التركيز على تعزيز قدرات الكشف من خلال المصنفات الذكية.

تتضمن المذكرة تحليلًا لخوارزميات التعلم الخاضع للإشراف المختلفة، المطبقة على مجموعة بيانات NSL-KDD، وهي معيار في مجال كشف التسلل. ومن أهم إسهامات هذا العمل تطوير حل هجين يتيح استخدام الشبكات العصبية التلافيفية أحادية البعد (DCNN1) ضمن بيئة WEKA. وقد تحقق هذا التكامل من خلال تصميم واجهة مخصصة تربط WEKA بمكتبات التعلم العميق القائمة على بايثون، متغلبًا على قيود WEKA المتأصلة في معالجة البيانات.

يسهّل الإطار المقترح تدريب النماذج وتقييمها بسلاسة من خلال الجمع بين واجهة WEKA سهلة الاستخدام وقوة التعلم العميق. يُمهّد هذا النهج الطريق لبناء أنظمة كشف تسلل أكثر متانة وتكيفًا. وقد يشمل العمل المستقبلي توسيع نطاق الحل ليشمل سيناريوهات كشف تسلل آني واختباره على مجموعات بيانات أكثر تنوعًا.

الكلمات المفتاحية: أنظمة كشف التسلل ، WEKA، التعلم الآلي، التعلم العميق، NSL-KDD، الشبكات العصبية التلافيفية.

Table des Matières