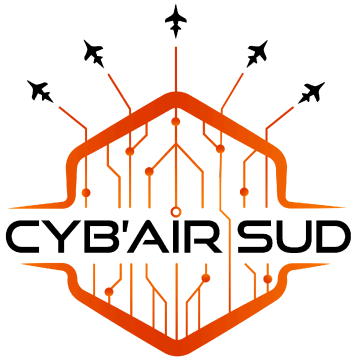


Qualifying Ambiguity: Technical and Ethical Challenges in Malware Classification

CFP 2025 — English Version

Speaker: Sylvio HOARAU

Presentation Objective

Highlight the technical and ethical challenges of automated malware classification in a context where some tools are ambiguous and classical and modern approaches (AI, threat intelligence) must coexist.

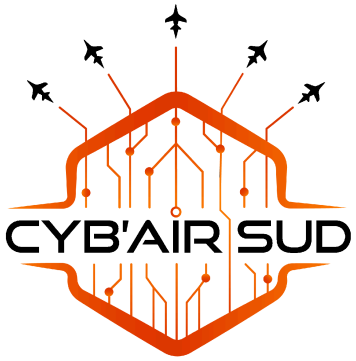
Abstract

In a context of proliferating cyber threats and increasing automation in detection, attribution, and classification, malware qualification represents a major challenge where technical and ethical concerns intersect. This presentation explores the inherent complexity of identifying threats in an ecosystem where tools can serve both legitimate and malicious purposes.

We first examine the fundamental paradigm balancing fast detection versus deep analysis, illustrating how some tools may be ambiguous and interpreted differently depending on the context of use. Comparative analysis between traditional methods (YARA signatures, IOCs) and AI-based approaches reveals a central dilemma: the former offer transparency but are vulnerable to evasion, while the latter promise robustness at the cost of deployment complexity.

Evaluating the effectiveness of detection systems raises critical questions regarding false positives and false negatives, especially in environments where unwarranted alerts can disrupt operations.

We will discuss how certain AI-based tools provide explainability, offering concrete ways to overcome the "black-box" nature of traditional models. This approach aims to build an ethical and contextual methodology for threat qualification, where technology assists but does not replace human judgment, acknowledging that ambiguity is not a bug to fix but a reality to manage.



Qualifier l'ambigu : défis techniques et éthiques de qualification de malware

CFP 2025 — Version Française

Intervenant : Sylvio HOARAU

Objectif de la présentation

Montrer les défis techniques et éthiques liés à la qualification automatique des malwares dans un contexte où certains outils sont ambigus, et où les approches classiques et modernes (IA, threat intelligence) doivent coexister.

Résumé

Dans un contexte de prolifération des cybermenaces et d'automatisation croissante de la détection, l'attribution et la qualification des malwares représentent un défi majeur où convergent enjeux techniques et considérations éthiques. Cette présentation explore la complexité inhérente à l'identification des menaces dans un écosystème où les outils peuvent servir des fins légitimes comme malveillantes.

Nous examinerons d'abord le paradigme fondamental opposant détection rapide et analyse approfondie, illustrant comment certains outils peuvent être ambigus et donc perçus différemment selon leur contexte d'utilisation. L'analyse comparative entre méthodes traditionnelles (signatures YARA, IOC) et approches basées sur l'intelligence artificielle révèle un dilemme central : les premières offrent transparence mais vulnérabilité au contournement, tandis que les secondes promettent robustesse au prix du coût de déploiement.

L'évaluation de l'efficacité des systèmes de détection soulève des questions critiques sur l'équilibre entre faux positifs et faux négatifs, particulièrement dans des environnements où une alerte injustifiée peut paralyser les opérations.

Nous aborderons comment l'utilisation de certains outils basés sur l'IA permet d'apporter de l'expliquabilité, offrant ainsi des pistes concrètes pour dépasser le caractère "boîte noire" des modèles traditionnels. Cette approche vise à construire une méthode éthique et contextuelle de la qualification des menaces, où la technologie facilite sans remplacer le ju-

gement humain, reconnaissant que l'ambiguïté n'est pas un bug à corriger mais une réalité à apprivoiser.