

الجمهورية الشعبية الديمقراطية الجزائرية
People's Democratic Republic of Algeria
وزارة التعليم العالي و البحث العلمي
Ministry of Higher Education and Scientific Research
المدرسة العليا للإعلام الآلي 8 ماي 5491 - سيدي بلعباس
Higher School of Computer Science
8 Mai 1945 - Sidi Bel Abbès



Graduation Thesis

To obtain the diploma of **Engineering's Degree**

Field of Study: **Computer Science**

Specialization: **Computer Systems Engineering (ISI)**

Theme

**AI based WAZUH alert classifier and analyer using
LLM**

Presented by
CHERRAK Ismail Anis

Defended on: **June, 2026**
In front of the jury composed of

Dr. Nemdili Amel
Dr. Baba-Ahmed Manel
Dr. Naib Rabia

President of the Jury
Thesis Supervisor
Examiner

Academic Year: 2025/2026

Abstract

The increasing volume and complexity of security alerts generated by modern SIEM systems pose a significant challenge for Security Operations Centers (SOC), particularly in distinguishing true positives from false positives and prioritizing incident response. This project proposes an intelligent pipeline designed to enhance the analysis of Wazuh alerts by leveraging fine-tuned large language models (LLMs).

The proposed system is centered around a fine-tuned LLM that performs the primary classification of alerts as true positives or false positives based on their contextual and semantic content. Beyond binary classification, the model is further utilized to conduct in-depth analysis of confirmed alerts, including the identification of attack patterns and the generation of relevant incident response (IR) actions.

To improve robustness and handle uncertainty, the pipeline incorporates a secondary validation mechanism for ambiguous cases, enabling more reliable decision-making. This approach allows the system not only to reduce false positives but also to transform raw alert data into actionable intelligence for SOC analysts.

The pipeline is evaluated on labeled Wazuh alert datasets, demonstrating improved classification performance and enhanced analytical capabilities compared to traditional rule-based or machine learning approaches. This work highlights the effectiveness of LLM-driven architectures in SIEM environments and contributes toward more intelligent and automated security operations.

Keywords: Wazuh, SIEM, Large Language Models, False Positive Reduction, Alert Classification, SOC, Incident Response, Cybersecurity

Résumé

Le volume croissant et la complexité des alertes générées par les systèmes SIEM modernes représentent un défi majeur pour les centres d'opérations de sécurité (SOC), notamment pour distinguer les vrais positifs des faux positifs et prioriser les réponses aux incidents. Ce projet propose un pipeline intelligent visant à améliorer l'analyse des alertes Wazuh en s'appuyant sur des modèles de langage de grande taille (LLM) finement ajustés.

Le système proposé repose sur un LLM affiné qui assure la classification principale des alertes en vrais positifs ou faux positifs, en exploitant leur contenu contextuel et sémantique. Au-delà de cette classification binaire, le modèle est également utilisé pour effectuer une analyse approfondie des alertes confirmées, incluant l'identification des schémas d'attaque et la génération d'actions de réponse aux incidents.

Afin d'améliorer la robustesse et de gérer les cas incertains, le pipeline intègre un mécanisme de validation secondaire pour les alertes ambiguës, permettant une prise de décision plus fiable. Cette approche permet non seulement de réduire les faux positifs, mais aussi de transformer les données brutes en informations exploitables pour les analystes SOC.

Le système est évalué sur des jeux de données annotés d'alertes Wazuh, montrant une amélioration des performances de classification et des capacités d'analyse par rapport aux approches traditionnelles. Ce travail met en évidence l'efficacité des architectures basées sur les LLM dans les environnements SIEM et contribue à des opérations de sécurité plus intelligentes et automatisées.

Mots-clés : Wazuh, SIEM, Modèles de langage, Réduction des faux positifs, Classification des alertes, SOC, Réponse aux incidents, Cybersécurité

ملخص

يمثل الحجم المتزايد وتعقيد التنبيهات الناتجة عن أنظمة SIEM الحديثة تحدياً كبيراً لمراكز عمليات الأمن، (SOC) خاصة في التمييز بين التنبيهات الصحيحة والخطئة وتحديد أولويات الاستجابة للحوادث. يقترح هذا المشروع نظاماً ذكياً لتحسين تحليل تنبيهات Wazuh بالاعتماد على نماذج اللغة الكبيرة (LLM) المحسنة. يركز النظام على نموذج لغة كبير مُدرَّب خصيصاً لتصنيف التنبيهات إلى صحيحة أو خاطئة اعتماداً على محتواها السياقي والدلالي. بالإضافة إلى ذلك، يُستخدم النموذج لتحليل التنبيهات المؤكدة بعمق، من خلال تحديد أنماط الهجمات واقتراح إجراءات الاستجابة للحوادث. ولتحسين موثوقية النتائج، يتضمن النظام آلية تحقق إضافية للتعامل مع الحالات غير الواضحة، مما يسمح باتخاذ قرارات أكثر دقة. يساهم هذا النهج في تقليل الإيجابيات الكاذبة وتحويل التنبيهات إلى معلومات قابلة للاستغلال من قبل محليي SOC. تم تقييم النظام باستخدام بيانات حقيقية لتنبيهات Wazuh حيث أظهرت النتائج تحسناً في دقة التصنيف وتعزيزاً لقدرات التحليل مقارنة بالأساليب التقليدية. تبرز هذه الدراسة أهمية استخدام نماذج اللغة في تطوير أنظمة SIEM أكثر ذكاءً وكفاءة.

الكلمات المفتاحية: Wazuh، SIEM نماذج اللغة، تقليل الإيجابيات الكاذبة، تصنيف التنبيهات، SOC الاستجابة للحوادث، الأمن السيبراني