Pengembangan Sistem Deteksi Intrusi Jaringan Berbasis Machine Learning yang Terintegrasi dengan Wazuh dan Zeek = Development of a Machine Learning-Based Network Intrusion Detection System Integrated with Wazuh and Zeek

Leonardo Jeremy Pong Pare Munda, author

Abstrak

Penelitian ini mengembangkan sistem deteksi intrusi jaringan dengan mengintegrasikan Zeek, model machine learning, dan Wazuh sebagai platform Security Information and Event Management (SIEM). Data dikumpulkan secara mandiri melalui proses monitoring jaringan selama dua jam menggunakan Zeek pada empat skenario: normal, GoldenEye, Slowloris, dan SlowHTTP. Log Zeek dianalisis secara real-time menggunakan model machine learning yang telah dilatih sebelumnya. Model dijalankan melalui pipeline Python yang membaca conn.log secara terus-menerus, memprediksi label “benign” atau jenis serangan, dan menyimpan hasilnya dalam file JSON. File ini kemudian dikirim ke Wazuh Server untuk dianalisis dan ditampilkan sebagai alert. Seluruh log Zeek juga dikirim ke Wazuh untuk mendukung proses threat hunting. Hasil pengujian menunjukkan bahwa sistem mampu mendeteksi serangan DoS secara efektif dengan detection rate mencapai 98,9% pada GoldenEye, 96,8% pada Slowloris, dan 77,3% pada SlowHTTPTest, jauh lebih tinggi dibandingkan pendekatan rule-based Zeek script yang hanya mencapai 46,7%, 55,1%, dan 28,4%. False positive rate juga tetap rendah, dengan ratarata di bawah 2% pada semua skenario. Selain itu, mean time to detect (MTTD) sistem menunjukkan performa yang responsif dan bersaing dengan Zeek script. Temuan ini membuktikan bahwa integrasi SIEM berbasis machine learning efektif dalam mendeteksi ancaman siber secara real-time di lingkungan yang terkendali.

.....This research develops a network intrusion detection system by integrating Zeek, a machine learning model, and Wazuh as a Security Information and Event Management (SIEM) platform. Data was independently collected through two hours of network monitoring using Zeek under four scenarios: normal, GoldenEye, Slowloris, and SlowHTTP. The Zeek logs were analyzed in real-time using a pre-trained machine learning model. The model was executed through a Python-based pipeline that continuously reads the conn.log, predicts whether each connection is benign or a specific attack type, and stores the results in a JSON file. This file is then sent to the Wazuh Server for analysis and displayed as alerts. In addition, all Zeek logs are forwarded to Wazuh to support threat hunting activities. Evaluation results show that the system can effectively detect DoS attacks, achieving a detection rate of 98.9% for GoldenEye, 96.8% for Slowloris, and 77.3% for SlowHTTPTest. These results significantly outperform the rule-based Zeek scripts, which only reached 46.7%, 55.1%, and 28.4% respectively. The system also maintains a low false positive rate, averaging below 2% across all scenarios. Moreover, the system demonstrates a responsive and competitive mean time to detect (MTTD) compared to Zeek scripts. These findings confirm that a machine learning-based SIEM integration can effectively enhance real-time cyber threat detection in controlled environments.