NU
NIIT UNIVERSITY
THE UNIVERSITY OF THE FUTURE

Hybrid Host-Based Intrusion Detection System Using System Call Analysis

Aryan Vats, Nitesh Jha, Shikhar Sharma
Mentored by: Dr. Anand Kumar Mishra

Concept Note

Background:Traditional Host-Based Intrusion Detection Systems (HIDS) often fail to detect stealth attacks efficiently and struggle with scalability. An adaptive hybrid system utilizing system call analysis is essential for modern cybersecurity.

Objectives:To design a lightweight, high-accuracy HIDS by combining machine learning (ML) and deep learning (DL) methods, ensuring real-time performance and handling imbalanced datasets.

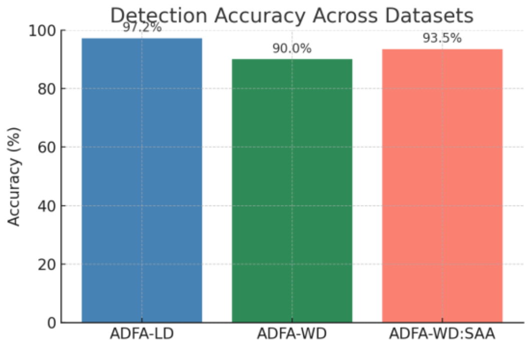
Methodology: ADFA-LD: TF-IDF (1–3 grams) + XGBoost,ADFA-WD: TF-IDF 5-grams on initial N system calls, stacking ensemble of KNN, RF, DT → XGBoost, AdaBoost, LightGBM.,ADFA-WD:SAA: CNN, LSTM, CNN+LSTM architectures with TF-IDF and SVD features.

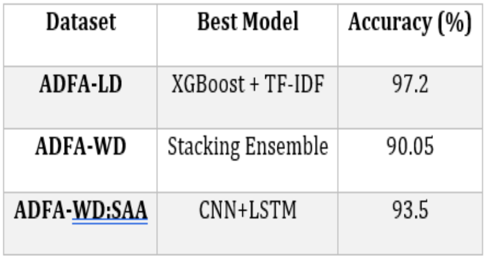
Datasets:ADFA-LD (Linux syscall traces), ADFA-WD (Windows DLL call traces), ADFA-WD:SAA (Windows stealth attack traces).

Key Results:ADFA-LD: 97.2% accuracy (TF-IDF + XGBoost),ADFA-WD: 90.05% accuracy and 94.78% recall with stacking ensemble,ADFA-WD:SAA: 93.5% accuracy with CNN+LSTM model, outperforming traditional ML models.

Significance:This work proposes a scalable, efficient HIDS framework suitable for real-time deployment in enterprise, government, and cloud systems.

Graphical Abstract





Dataset	Best Model	Accuracy (%)
ADFA-LD	XGBoost + TF-IDF	97.2
ADFA-WD	Stacking Ensemble	90.05
ADFA-WD:SAA	CNN+LSTM	93.5

Fig 2: Summary of Best Model Performance Across Datasets

Fig 1: Detection Accuracy Across Datasets