Ministry of ICT
ICT Research Institute
ITRC

In the name of God

Lateral Movement Attack Detection using Variational Autoencoders

Authors: Mostafa Shabani and Tala Tafazzoli*
ICT Security Faculty
Information and system security Group

Table of contents

Ministry of ICT
ICT Research Institute
ITRC

Ministry of ICT
ICT Research Institute
ITRC

01

Introduction

Ministry of ICT
ITRC ICT Research Institute

Introduction: The Stealthy Threat of Lateral Movement

Definition:
Lateral movement is a critical post-breach phase where an adversary pervasively compromises a network after an initial foothold.

Adversary's Goal:
To access high-value assets, culminating in large-scale data exfiltration or systemic service disruption.

The Core Challenge:
These attacks are designed to blend in with legitimate network traffic, thereby evading traditional signature-based security systems.

Limitations of Existing Methods:

- **Supervised Models:** Brittle when facing novel (zero-day) attack vectors due to their reliance on pre-labeled training data.
- **Unsupervised Models:** Prone to high false-positive rates and often fail to categorize specific attack typologies.

Ministry of ICT
ITRC ICT Research Institute

Problem Statement

How can we design an efficient and accurate detection system for lateral movement that provides:

High Accuracy: To reliably identify known attack patterns and minimize false alarms.
Computational Efficiency: To be fast enough for real-time monitoring and rapid threat response.
Generalization Potential: To be capable of flagging novel and unseen anomalies.

Our Proposed Solution: A Hybrid Deep Learning Framework based on a Variational Autoencoder

Ministry of ICT
ICT Research Institute
ITRC

02

Key Concepts

Key Concepts: Variational Autoencoder (VAE)

Generative Model:
The VAE is a neural network architecture that learns to model the underlying probability distribution of input data.

Core Components:
Encoder: Maps high-dimensional input data to a lower-dimensional, probabilistic latent space.
Decoder: Attempts to reconstruct the original input by sampling from this latent distribution.

Anomaly Detection Mechanism:
The VAE is trained exclusively on a robust baseline of normal system and network behavior.
Data points that the model fails to reconstruct accurately (i.e., those with a high reconstruction error) are identified as anomalies, signaling a deviation from the learned norm.

Ministry of ICT
ICT Research Institute
ITRC

03

Proposed Architecture

Key Concepts: Variational Autoencoder (VAE)

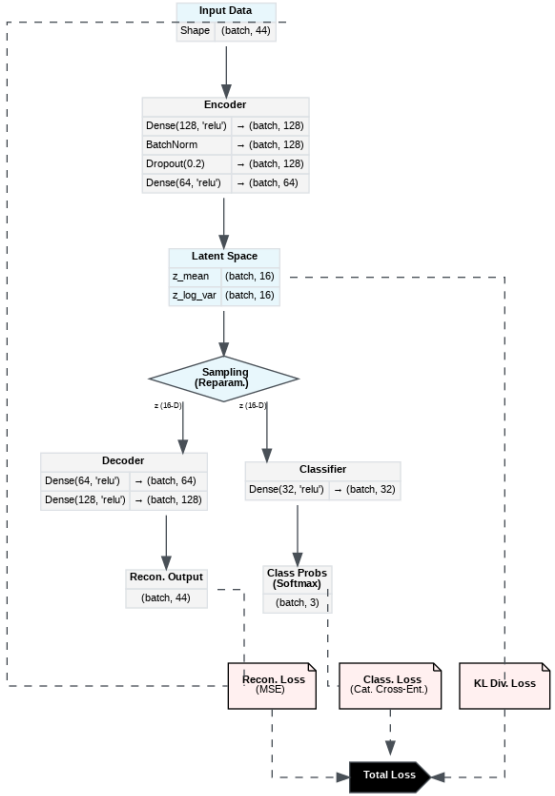
Dual-Purpose Design: A single, end-to-end model featuring a shared encoder with two distinct output heads.

1. **Reconstruction Head (Decoder):** For unsupervised, robust feature representation learning.
2. **Classification Head (Classifier):** For supervised, fine-grained threat classification.

Key Advantage: The reconstruction task acts as a powerful regularizer, compelling the encoder to learn a rich and generalizable latent representation that significantly benefits the classification task.

Data Flow:

1. A 44-dimensional feature vector enters the Encoder.
2. A 16-dimensional latent vector (z) is sampled.
3. This vector is simultaneously fed to the Decoder (for reconstruction) and the Classifier (for prediction).



Ministry of ICT
ITRC ICT Research Institute

Dataset and Feature Engineering

Dataset:

LMD-2023: The only benchmark corpus comprising Sysmon logs specifically for evaluating lateral movement detection methods.

1. Contains approximately 1.75 million log samples.
2. Highly imbalanced, with normal traffic constituting about 92% of the samples.
3. Three classes: 'Normal', 'EoRS' (Exploitation of Remote Services), and 'EoHT' (Exploitation of Hashing Techniques).

Feature Engineering Process:

Initial Extraction: Started with 93 features from raw Sysmon logs.

Feature Selection: Used PCA and model coefficient analysis to select 15 key conceptual features.

Encoding: Converted categorical features (e.g., hostnames) to numerical format using one-hot encoding.

Normalization: Scaled all numerical features to a standard [0, 1] range using Min-Max normalization.

Final Result: A 44-dimensional input feature vector for the model.

04

Experimental Setup

Ministry of ICT
ITRC ICT Research Institute

Experimental Setup

Environment: Google Colaboratory (Colab) with an NVIDIA Tesla T4 GPU and 16GB of VRAM.

Frameworks & Libraries: Python 3.10.12, TensorFlow 2.15.0 (with Keras), Scikit-learn 1.2.2.

Model Training Parameters:

- **Optimizer:** Adam with a learning rate of 0.001.
- **Batch Size:** 128.
- **Loss Function:** A composite function combining Reconstruction (MSE), KL Divergence, and Classification (Categorical Cross-Entropy) losses.

Overfitting Prevention: Employed EarlyStopping and ReduceLROnPlateau callbacks.

Evaluation Metrics: Accuracy, Precision, Recall, F1-Score, and AUC (Area Under the Curve).

Ministry of ICT
ICT Research Institute
ITRC

Experimental Setup

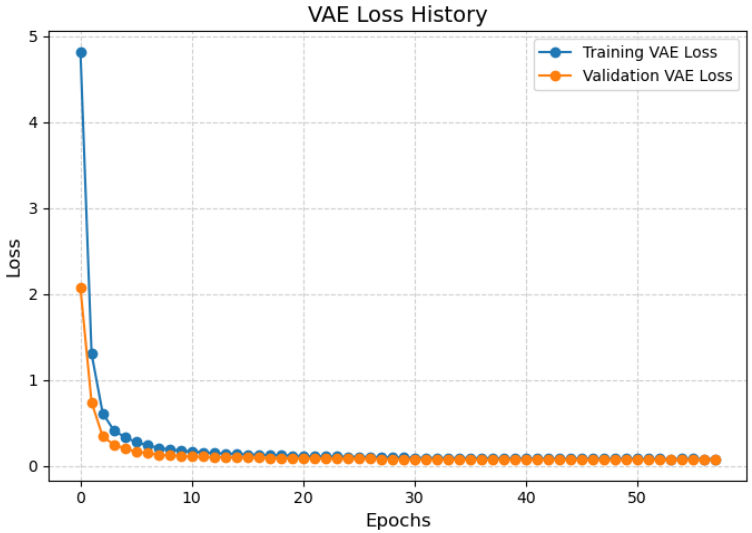
Model	Features	Cis	AUC	Prec	Recall	F1	Acc	Epochs	k-fold	T.E. Time
Lateral movement previous works										
MV (RF, LB, LoR)	4	2	-	-	-	0.66	99.62	N/A	✓	-
GRU DNN	8	2	-	93.23	-	-	96.68	60	✓	-
Ensemble ML	8	2	-	88.70	-	-	-	N/A	✓	-
SS DL	8	2	-	91.3	-	-	99.9	N/A	×	-
UML with JD	15	2	-	6	-	-	-	N/A	×	-
K-Means UML	27	2	81	-	-	-	-	N/A	×	-
RF	29	2	-	83.73	81.23	0.82	-	N/A	✓	00:00:02:06
LaBi	32	2	-	99.87	99.47	0.97	99.9	N/A	✓	00:00:11:28
RF	35	2	-	80.31	80.29	0.8	-	N/A	✓	00:00:03:11
ET	15	3	99.84	99.05	99.79	99.41	99.89	N/A	×	00:07:30:12
LSTM	15	3	95.82	95.11	94.36	95.55	98.93	30	×	00:15:44:18
This work. In each case, the best performers based on F1 score for LMD dataset										
VAE	44	3	99.6983	88.4756	88.5484	88.5029	98.3570	58	✓	00:00:02:54

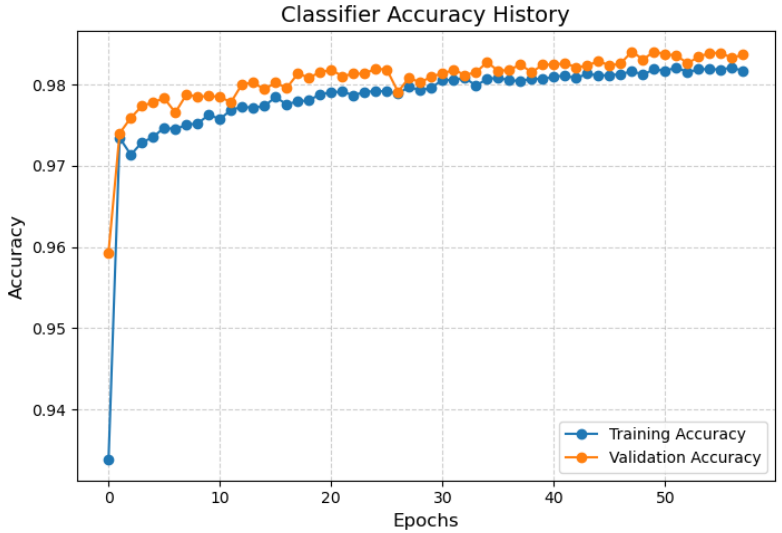
Key Findings:

Our model achieved an excellent AUC of 99.70%, demonstrating outstanding class separation capability. While the Extra Trees (ET) model shows a slightly higher F1-score, our VAE model is dramatically faster. A total execution time of 2.54 seconds versus 7.5 minutes for ET validates our model's viability for real-time, operational deployment.

Ministry of ICT
ICT Research Institute
ITRC

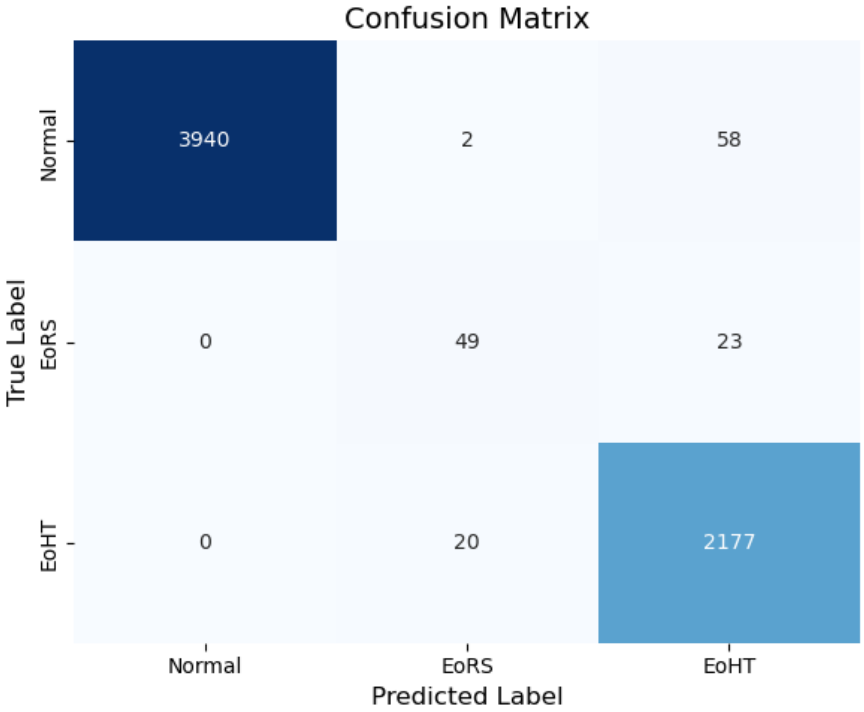
Model evaluation





Ministry of ICT
ICT Research Institute
ITRC

Confusion Matrix



05

Discussion and Analysis

Ministry of ICT
ICT Research Institute
ITRC

Conclusion & Future Work

Key Contributions of This Work:

1. A Novel Hybrid Framework: We proposed and successfully implemented a VAE-centric hybrid framework, proving the viability of generative models for this complex cybersecurity challenge.

2. A Compelling Balance of Accuracy and Efficiency: Our empirical results demonstrate a strong balance between high discriminative power (AUC of 99.70%) and operational speed, establishing the VAE as a powerful and viable tool against Advanced Persistent Threats (APTs).

Future Work:

1. Enhancing Interpretability: Analyzing the learned latent space to better understand the model's decision-making process.
2. Exploring Unsupervised Capabilities: Formally leveraging reconstruction error as a standalone mechanism for zero-day threat detection.
3. Hybridization with Other Methods: Combining VAEs with Graph Neural Networks (GNNs) or Bayesian approaches to enhance detection capabilities.
4. Evaluating on more diverse and noisy real-world datasets to further test model robustness.

Thank you for your attention.

Do you have any questions?