SSH Shell Attacks

ANDREA BOTTICELLA*, Politecnico di Torino, Italy
ELIA INNOCENTI*, Politecnico di Torino, Italy
RENATO MIGNONE*, Politecnico di Torino, Italy
SIMONE ROMANO*, Politecnico di Torino, Italy

This paper introduces a comprehensive machine learning framework to analyze SSH shell attack sessions, leveraging both supervised and unsupervised learning techniques. Using a dataset of 230,000 unique Unix shell attack sessions, the framework aims to classify attacker tactics based on the MITRE ATT&CK framework and uncover latent patterns through clustering. The key contributions of this work are:

- Development of a robust pre-processing pipeline to analyze temporal trends, extract numerical features, and evaluate intent distributions from large-scale SSH attack session data.
- Implementation of supervised classification models to accurately predict multiple attacker tactics, supported by hyperparameter tuning and feature engineering for enhanced performance.
- Application of unsupervised clustering techniques to uncover hidden patterns in attack behaviors, leveraging visualization tools and cluster analysis for fine-grained categorization.
- Exploration of advanced language models, such as BERT and Doc2Vec, for representation learning and fine-tuning to improve intent classification and session interpretation.

CCS Concepts: • **Security and privacy** → **Intrusion detection systems**; **Malware and its mitigation**; • **Computing methodologies** → *Supervised learning by classification*; *Unsupervised learning*; *Natural language processing*.

Additional Key Words and Phrases: SSH shell attacks, machine learning, supervised learning, unsupervised learning, language models, security logs, intrusion detection.

Contents

1 INTRODUCTION

1.1 Motivation

Security logs play a crucial role in understanding and mitigating cyber attacks, particularly in the domain of network and system security. With the increasing sophistication of cyber threats, analyzing and interpreting

*All authors contributed equally to this research.

Authors' Contact Information: Andrea Botticella, andrea.botticella@studenti.polito.it, Politecnico di Torino, Turin, Italy; Elia Innocenti, elia.innocenti@studenti.polito.it, Politecnico di Torino, Turin, Italy; Renato Mignone, renato.mignone@studenti.polito.it, Politecnico di Torino, Turin, Italy; Simone Romano, simone.romano@studenti.polito.it, Politecnico di Torino, Turin, Italy.

Publication date: December 2024.

security logs has become paramount for detecting, preventing, and responding to potential security breaches. Unix shell attacks, especially those executed through SSH, represent a significant vector for potential system compromises.

The complexity of security log analysis stems from several key challenges:

- Logs are often unstructured and contain ambiguous or malformed text.
- Manual parsing and interpretation of logs is time-consuming and error-prone.
- The sheer volume of log data makes comprehensive manual review impractical.

These challenges underscore the need for automated, intelligent approaches to log analysis that can efficiently extract meaningful insights and identify potential security threats.

1.2 Objective

The primary objective of this research is to develop and evaluate machine learning techniques for automatic analysis and classification of SSH shell attack logs. Specifically, we aim to:

- Explore and preprocess a large dataset of Unix shell attack sessions.
- Apply supervised learning techniques to classify attack tactics based on session characteristics.
- Utilize unsupervised learning methods to discover patterns and clusters in attack sessions.
- Investigate the potential of advanced language models in understanding and categorizing attack intents.

By leveraging machine learning approaches, we seek to:

- Automate the process of log analysis and intent classification.
- Provide security professionals with insights into attack strategies.
- Develop a framework for understanding and categorizing SSH shell attacks using the MITRE ATT&CK tactics as a reference.

The significance of this research lies in its potential to enhance cybersecurity threat detection and response capabilities by transforming complex, unstructured log data into actionable intelligence.

2 BACKGROUND

2.1 Security Logs and Attack Analysis

Security logs represent a critical source of information for understanding system vulnerabilities and potential cyber attacks. These logs capture detailed records of system events, network interactions, and user activities, providing valuable insights into potential security breaches.

In the context of SSH shell attacks, logs document the sequence of commands executed during a malicious session, enabling security researchers to analyze attacker behaviors, techniques, and potential system impacts. However, the manual analysis of these logs is challenging due to their volume, complexity, and often non-standard formatting.

2.2 MITRE ATT&CK Framework

The MITRE ATT&CK (Adversarial Tactics, Techniques, and Common Knowledge) framework provides a comprehensive knowledge base of adversary tactics and techniques observed in real-world cyber attacks. This framework serves as a standardized methodology for understanding and categorizing attack strategies.

For our research, we focus on seven key intents derived from the MITRE ATT&CK framework:

- **Persistence:** Techniques used by adversaries to maintain system access across restarts or credential changes.
- **Discovery:** Methods for gathering information about the target system and network environment.
- **Defense Evasion:** Strategies to avoid detection by security mechanisms.

- **Execution:** Techniques for running malicious code on target systems.
- **Impact:** Actions aimed at manipulating, interrupting, or destroying systems and data.
- **Other:** Less common tactics including Reconnaissance, Resource Development, Initial Access, etc.
- **Harmless:** Non-malicious code or actions.

2.3 Dataset Overview

Our research utilizes a comprehensive dataset of SSH shell attacks collected from a honeypot deployment. Key characteristics of the dataset include:

- Approximately 230,000 unique Unix shell attack sessions
- Recorded after SSH login
- Stored in Parquet format for efficient data processing
- Columns include:
  - Session ID
  - Full session text (command sequences)
  - Timestamp
  - Intent labels based on MITRE ATT&CK tactics

It is important to note that the dataset's labels are automatically generated through a research project and may contain potential classification errors. This inherent uncertainty adds an additional layer of complexity to our analysis and highlights the importance of robust machine learning techniques.

2.4 Research Approach

Our research employs a multi-faceted approach to SSH shell attack log analysis:

- Comprehensive data exploration and preprocessing
- Supervised learning for attack intent classification
- Unsupervised learning for attack pattern discovery
- Advanced language model exploration

By combining these techniques, we aim to develop a comprehensive framework for understanding and categorizing SSH shell attacks, ultimately contributing to improved cybersecurity threat detection and response strategies.

3 DATA EXPLORATION AND PRE-PROCESSING

This section …

3.1 Introduction

Brief introduction to the data exploration and pre-processing tasks.

…

3.2 Dataset Preparation

Loading the dataset and initial inspection.

…

3.3 Temporal Analysis

Analysis of when the attacks were performed.

…

3.4 Feature Extraction

Extracting features from the attack sessions.

…

3.5 Common Words Analysis

Identifying the most common words in the sessions.

…

3.6 Intent Distribution

Analyzing the distribution of intents over time.

…

3.7 Text Representation

Converting text into numerical representations (BoW, TF-IDF).

…

4 SUPERVISED LEARNING - CLUSTERING

This section …

4.1 Introduction

Overview of the supervised learning task and its objectives.

…

4.2 Data Splitting

Splitting the dataset into training and test sets.

…

4.3 Baseline Model Implementation

Implementing and evaluating baseline models.

…

4.4 Hyperparameter Tuning

Tuning hyperparameters and evaluating performance.

…

4.5 Result Analysis

Analyzing the results for each intent.

…

4.6 Feature Experimentation

Exploring different feature combinations and their impact on performance.

…

5 UNSUPERVISED LEARNING - CLUSTERING

This section …

Publication date: December 2024.

5.1 Introduction

Overview of the clustering task and its objectives.

…

5.2 Determine the Number of Clusters

Using methods like the elbow method or silhouette analysis.

…

5.3 Hyperparameter Tuning

Tuning other hyperparameters, if any.

…

5.4 Cluster Visualization

Visualizing the clusters through t-SNE.

…

5.5 Cluster Analysis

Analyzing the characteristics of each cluster.

…

5.6 Intent Homogeneity

Assessing if clusters reflect intent division.

…

5.7 Specific Attack Categories

Associating clusters with specific attack categories.

…

6 LANGUAGE MODEL EXPLORATION

This section …

6.1 Introduction

Overview of the language models task and its objectives.

…

6.2 Pretraining

Pretraining Doc2Vec or using a pretrained Bert model.

…

6.3 Model Fine-tuning

Fine-tuning the last layer of the network.

…

6.4 Learning Curves

Plotting learning curves and determining the optimal number of epochs.

…

7 CONCLUSION

This section …

7.1 Subsection 1

…

7.2 Subsection 2

…

Publication date: December 2024.