PCAP-Based Intrusion Detection Using a Local LLM

Project Overview

This project analyses network traffic captured in PCAP files to identify potential cyber attacks. It focuses on detecting three specific attack types:

- SYN Flood
- Port Scanning
- DNS Exfiltration

The system combines traditional packet analysis and traffic aggregation with a locally hosted Large Language Model (LLM). The LLM is used to classify and explain suspicious network behavior based on summarized traffic patterns. All processing is performed locally, without reliance on cloud-based APIs, ensuring data privacy.

Key Features

- Parses PCAP files using packet inspection
- Aggregates traffic by IP addresses, ports, DNS queries, and time
- Detects SYN floods, port scans, and DNS exfiltration
- Uses a local LLM for attack classification and reasoning
- Produces human-readable explanations for detected attacks

Attack Detection Approaches

SYN Flood Detection

- Aggregates TCP traffic by source and destination IP
- Tracks the number of TCP SYN and ACK packets
- Summarizes traffic duration and packet imbalance

- LLM determines whether the behavior matches a SYN flood

Port Scan Detection

- Aggregates TCP traffic by source IP
- Tracks the number of unique destination ports contacted
- Summarizes scanning behavior over time
- LLM determines whether the behavior matches a port scan

DNS Exfiltration Detection

- Aggregates DNS queries by source IP
- Tracks query frequency and query length
- Highlights unusually long or frequent DNS queries
- LLM determines whether the behavior matches DNS exfiltration

Role of the Local LLM

The LLM is used as a semantic classification and explanation layer, not as a low-level packet classifier. Specifically, the LLM:

- Receives summarized traffic descriptions
- Classifies traffic as one of: SYN Flood, Port Scan, DNS Exfiltration, Benign
- Generates natural-language explanations describing its reasoning

Domain Knowledge

Domain knowledge expertise was used to come with reliable feature extraction methods and to evaluate and classify our summarizations. Domain expertise is the result of the laboratory work we had, which included:

- Study common patterns associated with SYN floods, port scans, and DNS exfiltration

- Select relevant packet and flow-level features
- Validate aggregation logic and detection behavior
- Inform LLM prompt design

The dataset was not used to fine-tune the LLM. Detection logic is based on aggregation, while the LLM provides reasoning and classification.

System Architecture

PCAP File → Packet Parsing (Scapy) → Traffic Aggregation (IPs, Ports, DNS, Time) → Traffic Summarization (Text) → Local LLM Classification → Attack Alerts + Explanations

Requirements

- Python 3.9+
- Jupyter Notebook or local Python environment
- Scapy
- Ollama (or another local LLM runner)

Python Dependencies

- pip install scapy
- Local LLM Setup (Example: Ollama)
- ollama pull mistral

Usage

1. Place a PCAP file in the project directory
2. Load the PCAP file in the notebook or script
3. Run the detection pipeline functions
4. Review LLM-generated alerts and explanations

Output

The system outputs a list of detected traffic patterns, each including:

- Attack type (or benign)
- Confidence score
- Explanation generated by the LLM

Design Rationale

- Traditional packet analysis is efficient and reliable for detection
- LLMs excel at reasoning and explanation, not raw packet classification
- Separating detection and explanation improves accuracy and transparency
- Running the LLM locally preserves privacy and security

Limitations

- Thresholds and aggregation windows are static
- Performance depends on the quality of packet captures
- LLM output may vary slightly between runs
- Not designed for real-time deployment

Conclusion

This project demonstrates a hybrid intrusion detection approach that combines traditional network analysis with modern language models. By using a local LLM for classification and explanation, the system provides interpretable and privacy-preserving detection of SYN floods, port scans, and DNS exfiltration.