DDoS Detection Using Machine Learning

Andrea Wynn, Bohdan Vakhitov, Duncan McKee

Advised by Dr. Michael Wollowski and MIT Lincoln Lab

ROSE-HULMAN INSTITUTE OF TECHNOLOGY

Background

- Computer networks vital to transferring important information and function of software systems
- Integrity may be compromised by malicious actors

Introduction

- Study random forests as a tool for DDOS attack detection
- Detect malicious activity on a network
- Perform a meta-analysis of current work on DDOS attack detection (ML vs rule-based methods)
- Evaluate the effectiveness of ML and classical AI methods in detecting malicious activities (i.e DDoS attacks)
- Test accuracy of detection methods

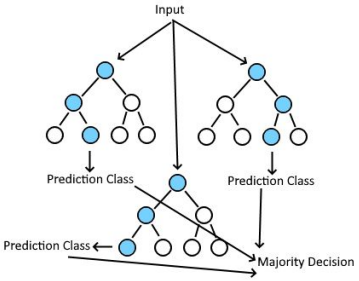
DDoS Detection

- DDoS attacks overload a network
- Disrupt normal network behavior
- Multiple methods to detect attacks:
  - Pre programmed thresholds
  - Random Forest Classifier
  - Feed-Forward Network
  - Other ML methods

Data

- Individual packet flows (forward/reverse)
- Each dataset entry contains 84 features
- Each packet is either part of benign or DDoS traffic
- Total of 12,794,627 records

Random Forest Classifier Model



- Majority based decision
- Classifies input data

- 70/30 testing split
- 200 estimators
- Maximum depth of 5

Ablation Study Results

- Validating high accuracy achieved with ML models
- Identifying and selecting most important features
- Removing some features, re-training, finding resulting accuracy
- Most important feature: Minimum size observed forward

Features Kept	Accuracy	Most Important Feature
All	0.9931	Fwd Seg Size Min
Bottom 50%	0.9622	Pkt Size Avg
Top 50%	0.9930	Fwd Seg Size Min
Top 1	0.9127	Fwd Seg Size Min
Bottom 5	0.5762	Active Mean

Meta Analysis

Other paper results

- Probability based models
- Information theory based models
- Distance measurement based models
- Machine learning models
- Signal processing based models

Open source Industry tools

- Most well known tools
  - OSSEC
  - Snort
- Log based intrusion detection
- Packet sniffer and rule based detection
- Some presence of machine learning models
- Mainly hand trained rule models

Proprietary DDOS protection technologies

- Google Cloud Armor
  - Rule-Based system
  - Utilizes ML architecture for attack detection
  - Application layer threat protection

Conclusion

- ML is a promising potential approach, because it can adapt to the rapidly changing field of cyberattack detection
- Future work could include combining ML with existing detection tools, ensemble methods, identifying multiple attacks types
- Handcrafted detection methods are still popular due to their explainability and ease of implementation