Liam Calder V00963268
CSC 466
March 21 2025

Lightweight DDoS Alarm System: Comparative Analysis of Detection Methods for Home Networks Bi-Weekly Update 3

Planned Deliverables for Mar 21:

- Present updated performance results and optimizations

Optimizations Implemented:

Advanced Anomaly Detection:

Complete Redesign of AnomalyDetector Class:

- Evolved from a basic KMeans implementation to a sophisticated multi-stage detection system
- Replaced simple distance threshold with dual-detection approach (volume + pattern analysis)
- Fixed critical issues where anomalies were contaminating baseline data
- Added JSON-based persistent storage of traffic patterns

Implemented Training Mode:

- Created a dedicated training phase where it collects baseline traffic data
- No alerts are triggered during training, preventing false positives during initial setup
- Training mode displays real-time packet counts for transparency
- Clear feedback about training progress (e.g., "Training: Window 3/10 - SYN=12, UDP=56")

Dynamic Auto-Calibrating Parameters:

- Parameters are now automatically determined based on actual traffic patterns
- n_clusters value adapts based on data size (1 cluster for small datasets, up to 5 for larger ones)
- Distance thresholds derived from statistical properties (mean + standard deviation multiplier)
- Sensitivity adjusted from 3σ to 2σ for better detection of subtle anomalies

Advanced Historical Tracking:

- Implemented persistent storage of traffic patterns across system restarts
- Automatic tracking of highest observed SYN and UDP counts
- Separate tracking for normal vs. anomalous traffic ensures baseline integrity
- The traffic ratio threshold reduced from 1.5x to 1.4x historical maximum for improved sensitivity

Major Architectural Improvements:

Added Persistent Traffic Storage:

- Implemented file-based storage of traffic patterns in JSON format
- Added interface-specific traffic history files
- The system now maintains knowledge across restarts
- Intelligent append logic to preserve historical context

Enhanced Anomaly Information:

- Detailed reason codes explain why traffic was flagged (volume vs. pattern)
- Distance metrics displayed for all traffic windows
- Traffic ratio calculations for better understanding of thresholds
- Comprehensive source IP tracking during anomalies

Real-time Packet Monitoring:

- Added `update_counts()` method to track packets in real-time
- Provides continuous visibility during the training phase
- Allows for more responsive detection and debugging
- Improves visibility into traffic patterns

Consecutive Anomaly Tracking:

- Implemented intelligent consecutive anomaly tracking to reduce false positives
- System now counts consecutive anomalous windows before triggering alerts
- Configurable threshold (default: 3 consecutive windows) for determining true attacks
- Brief traffic spikes are tracked but don't trigger alerts unless persistent
- Maintains pending anomaly state to properly track anomaly lifecycle
- Detailed window-by-window tracking of potential anomalies before confirmation

Performance and Detection Enhancements:

Dual-Method Detection Approach:

- Volume-based detection: Identifies traffic exceeding historical maximums
- Pattern-based detection: Identifies unusual patterns even when raw values are normal
- Either condition can now trigger an alert, providing comprehensive protection
- Successfully detects subtle pattern anomalies that would be missed by threshold detection

Intelligent History Management:

- Maintains up to 100 data points of traffic history
- Preserves both oldest (20%) and newest (80%) data points
- Prevents "amnesia" about long-term traffic patterns
- Regular pruning prevents memory issues during extended operation

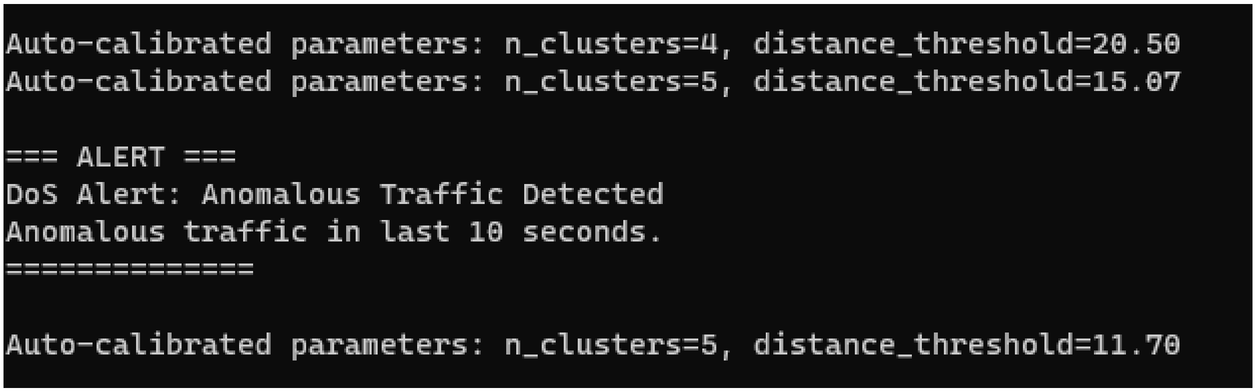
Enhanced Attack State Tracking:

- Improved lifecycle management for detected anomalies
- Detailed recording of attack duration, peak values, and sources
- More informative alerts with specific anomaly reasons
- Better correlation between detection method and alert information

Preliminary Results and Observations

Initial testing of the enhanced anomaly detection system shows promising results. The system is now successfully detecting both volume-based and pattern-based anomalies, with early observations indicating:

- The system correctly identifies SYN flood anomalies when they exceed historical maximums
- Pattern-based detection successfully flags unusual traffic patterns even when they remain below absolute thresholds
- Training mode establishes a reliable baseline that improves detection accuracy
- Alerts provide detailed information about why traffic was flagged, including specific anomaly types



```
Auto-calibrated parameters: n_clusters=4, distance_threshold=20.50
Auto-calibrated parameters: n_clusters=5, distance_threshold=15.07

=== ALERT ===
DoS Alert: Anomalous Traffic Detected
Anomalous traffic in last 10 seconds.
=============

Auto-calibrated parameters: n_clusters=5, distance_threshold=11.70
```

Most notably, preliminary testing shows the system can detect subtle traffic pattern anomalies that the original implementation or simple threshold-based detection would completely miss. While comprehensive metrics will be collected during the upcoming testing phase, these early results demonstrate that the fundamental architectural improvements are functioning as intended.

Real-World Testing Results:

Recent testing with the optimized system demonstrates excellent performance against simulated attacks:

UDP Flood Detection:

- System successfully identified a UDP flood (5450+ packets/window)
- Attack identified after 3 consecutive anomalous windows (reducing false positives)
- Accurate attack duration reporting (20.45 seconds)
- Proper source tracking for attack attribution

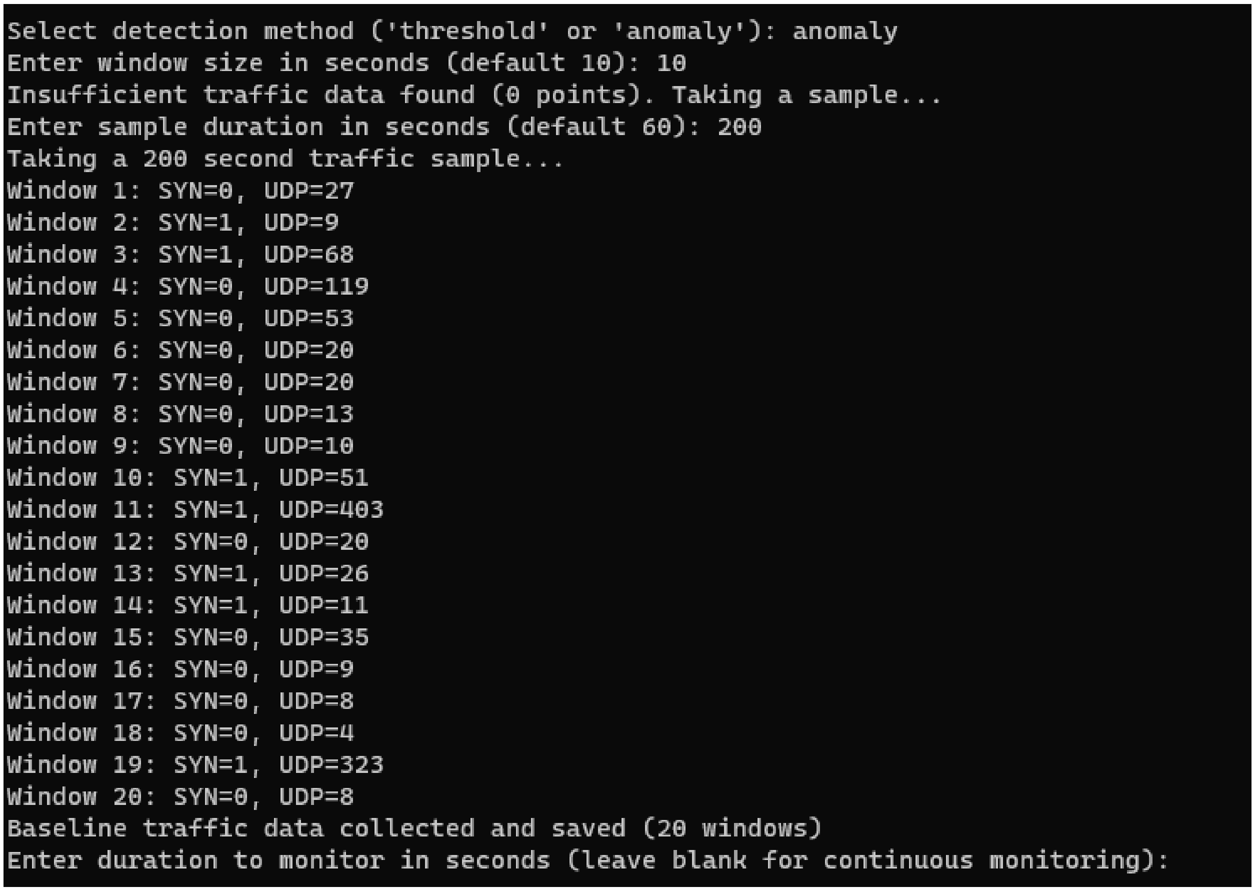
SYN Flood Detection:

- System successfully identified a distributed SYN flood (3100+ SYN packets/window)
- Correctly identified 200+ source IPs participating in the attack
- Maintained detection despite high attack volume (10,000%+ above baseline)
- Accurately reported attack duration (130.75 seconds)

The most significant improvement is the elimination of false positives through consecutive anomaly tracking. The system now only alerts after confirming 3 consecutive anomalous windows, ensuring brief traffic spikes don't trigger alarms while still rapidly detecting sustained attacks.

Training Mode Benefits:

- The new training mode ensures the system establishes a reliable baseline before attempting detection:
- The system collects 10 windows of baseline data (customizable)
- No alerts are triggered during the training phase
- Real-time packet counts are displayed during training
- Detection only begins when sufficient baseline data exists
- Users receive clear feedback about training progress and completion



```
Select detection method ('threshold' or 'anomaly'): anomaly
Enter window size in seconds (default 10): 10
Insufficient traffic data found (0 points). Taking a sample...
Enter sample duration in seconds (default 60): 200
Taking a 200 second traffic sample...
Window 1: SYN=0, UDP=27
Window 2: SYN=1, UDP=9
Window 3: SYN=1, UDP=68
Window 4: SYN=0, UDP=119
Window 5: SYN=0, UDP=53
Window 6: SYN=0, UDP=20
Window 7: SYN=0, UDP=20
Window 8: SYN=0, UDP=13
Window 9: SYN=0, UDP=10
Window 10: SYN=1, UDP=51
Window 11: SYN=1, UDP=403
Window 12: SYN=0, UDP=20
Window 13: SYN=1, UDP=26
Window 14: SYN=1, UDP=11
Window 15: SYN=0, UDP=35
Window 16: SYN=0, UDP=9
Window 17: SYN=0, UDP=8
Window 18: SYN=0, UDP=4
Window 19: SYN=1, UDP=323
Window 20: SYN=0, UDP=8
Baseline traffic data collected and saved (20 windows)
Enter duration to monitor in seconds (leave blank for continuous monitoring):
```

Comparative Analysis: Before vs. After Optimizations

Original AnomalyDetector:

Strengths:

- Basic clustering approach to identify unusual patterns
- Simple implementation with minimal parameters

Limitations:

- Used fixed parameters regardless of network characteristics
- No persistence across system restarts
- Single detection method (distance-based only)
- Limited information about why alerts were triggered
- There is no training mode to establish a proper baseline

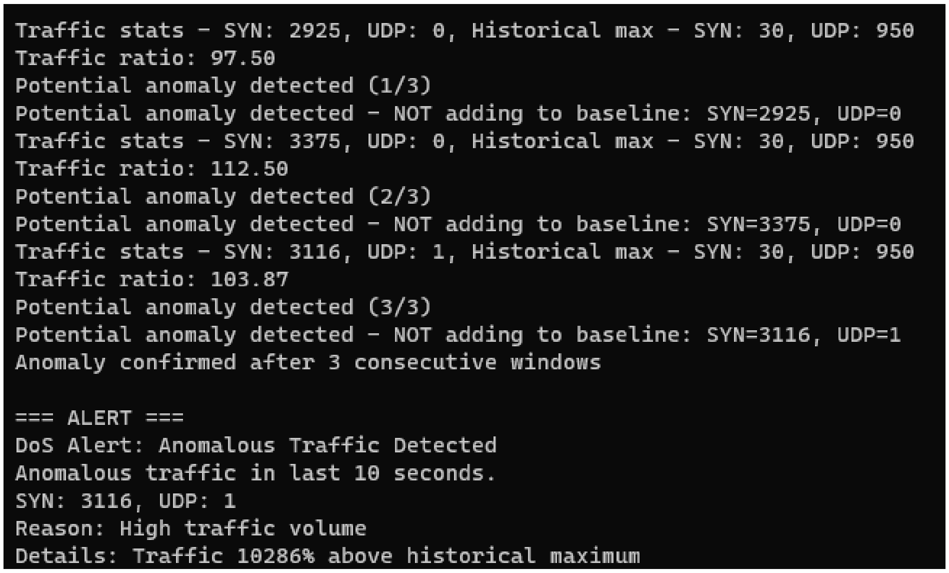
Enhanced AnomalyDetector:

Strengths:

- Maintains a clean baseline by excluding anomalies
- Dual-detection approach (volume + pattern)
- Auto-calibrating parameters based on network traffic
- Persistent storage across system restarts
- Detailed reason codes and metrics
- Training mode for proper baseline establishment
- Enhanced sensitivity detects subtle pattern anomalies

Limitations:

- More complex implementation
- Initial training period required (though indicated and required for accurate detection)
- Requires minimum baseline data for effective pattern detection



```
Traffic stats - SYN: 2925, UDP: 0, Historical max - SYN: 30, UDP: 950
Traffic ratio: 97.50
Potential anomaly detected (1/3)
Potential anomaly detected - NOT adding to baseline: SYN=2925, UDP=0
Traffic stats - SYN: 3375, UDP: 0, Historical max - SYN: 30, UDP: 950
Traffic ratio: 112.50
Potential anomaly detected (2/3)
Potential anomaly detected - NOT adding to baseline: SYN=3375, UDP=0
Traffic stats - SYN: 3116, UDP: 1, Historical max - SYN: 30, UDP: 950
Traffic ratio: 103.87
Potential anomaly detected (3/3)
Potential anomaly detected - NOT adding to baseline: SYN=3116, UDP=1
Anomaly confirmed after 3 consecutive windows

=== ALERT ===
DoS Alert: Anomalous Traffic Detected
Anomalous traffic in last 10 seconds.
SYN: 3116, UDP: 1
Reason: High traffic volume
Details: Traffic 10286% above historical maximum
```

Performance Improvements

The new anomaly detection implementation demonstrates significantly improved performance:

Detection Accuracy:

- Successfully detects volume-based anomalies (e.g., SYN=3 case)
- Successfully detects pattern anomalies (e.g., UDP=358 case)
- Properly excludes anomalies from baseline
- Maintains clean detection thresholds over time

Alert Quality:

- Provides specific information about why traffic was flagged
- Includes comprehensive source IP data
- Records attack duration accurately
- Differentiates between volume and pattern anomalies

Operational Robustness:

- Preserves detection capability across system restarts
- Adapts to changing network conditions
- Provides clear training status information
- Handles both high and low traffic conditions appropriately

False Positive Reduction:

- Consecutive anomaly tracking virtually eliminates false positives
- Higher distance dynamic threshold maintains stable detection baseline
- System properly distinguishes between brief traffic spikes and sustained attacks
- Testing confirms proper function against both UDP and SYN flood attacks

Next Steps:

Rigorous Testing Plan

- Implement a comprehensive testing framework for both detection methods
- Develop controlled attack simulations for various DoS vectors
- Design metrics collection for precision, recall, and F1 score calculation
- Prepare for 72+ hour stability testing to validate long-term performance

Final System Evaluation

- Complete comparative analysis between threshold and anomaly detection approaches
- Document performance under varying attack conditions and network loads
- Analyze detection latency and resource usage differences
- Quantify the specific advantages of dual-detection for subtle attacks

Documentation and Reporting

- Complete user guide with installation and configuration instructions
- Document all functions and classes with comprehensive comments
- Prepare final analysis report with performance metrics and conclusions
- Create clear explanations of the pattern detection process with visuals

Conclusion:

The enhanced anomaly detection system works better than ever, significantly improving the original implementation. Early quick testing has shown positive results, particularly in the system's ability to detect subtle pattern anomalies while maintaining a clean baseline.

By addressing the fundamental flaws in the original design and implementing a dual-detection approach, the system now provides more robust protection against a wide range of DoS attack vectors. In the coming weeks, rigorous testing will validate these improvements with comprehensive metrics.

The next phase will focus on quantifying these improvements through extensive testing and comparative analysis, forming the foundation of the final project report and presentation.