vehicle-agnostic IDS

Haneen Almassri 2185044
Master's in International Cybersecurity and Cyber Intelligence

Overview

This document provides a concise summary of my project focused on developing a vehicle-agnostic Intrusion Detection System (IDS) for Controller Area Network (CAN) bus security in modern vehicles. The complete implementation—including all source code, dataset processing scripts, and detailed documentation—is available in the project repository.

Repository: https://github.com/Haneen961/vehicle-agnostic-IDS

1. Motivation and Problem Statement

Modern vehicles have evolved into complex cyber-physical systems with over 100 Electronic Control Units (ECUs) communicating via the CAN bus protocol. While CAN is efficient and reliable, it was designed for closed, trusted environments and lacks fundamental security mechanisms:

- No encryption – All messages are transmitted in plaintext
- No authentication – Any node can send messages without verification
- No message integrity – Messages can be altered without detection

This vulnerability has been demonstrated through numerous real-world attacks where researchers have gained unauthorized control over critical vehicle functions like braking, steering, and acceleration. While Intrusion Detection Systems (IDS) offer a promising defense, existing solutions face two critical limitations:

1. Vehicle-specific training: Most IDS models are trained on data from single vehicles, making them ineffective when deployed on different vehicle models
2. Data scarcity: Lack of high-quality, publicly available datasets with realistic attacks recorded under actual driving conditions

2. Core Innovation: Vehicle-Agnostic Approach

The key innovation of this project is developing an IDS that works across different vehicle platforms without requiring retraining for each specific vehicle model. This is achieved through:

Statistical Feature Abstraction

Instead of analyzing specific CAN IDs or message semantics (which vary between vehicles), the system focuses on fundamental statistical properties of network traffic that remain consistent across platforms:

- Message frequency distributions
- Timing characteristics (inter-arrival times)
- Identifier entropy and distribution patterns
- Network load characteristics

Multi-Source Dataset Creation

To train a truly vehicle-agnostic model, I created a comprehensive dataset by aggregating and normalizing benign data from four major public CAN datasets:

- CAN-MIRGU (modern autonomous-capable vehicle)
- Car hacking dataset for the intrusion detection
- Car hacking attack and defence challenge
- OTIDS Dataset

This multi-source approach provides diverse examples of "normal" CAN behavior across different vehicle platforms, forcing the model to learn generalized patterns rather than vehicle-specific quirks.

3. Technical Implementation

Architecture: Hybrid Anomaly Detection

The system employs a two-stage hybrid architecture combining deep learning and traditional machine learning:

CAN Traffic → Feature Extraction → [Autoencoder + One-Class SVM] → Combined Decision

Stage 1: Deep Auto Encoder

- Purpose: Learn compressed representations of normal traffic patterns

- Training: Exclusive use of benign multi-vehicle data
- Detection: High reconstruction error indicates anomalies
- Advantage: Captures complex, non-linear patterns in high-dimensional data

Stage 2: One-Class Support Vector Machine (OCSVM)

- Purpose: Define boundaries of normal behavior in latent space
- Training: Uses autoencoder's latent representations
- Detection: Points outside boundary classified as anomalies
- Advantage: Provides statistical measure of "normalcy"

Combined Decision Logic

- Uses AND operation: Flag as attack only if BOTH models agree
- Benefit: Reduces false positives while maintaining high detection rates

Key Technical Challenges and Solutions

Challenge	Solution Implemented
Vehicle-specific patterns	Statistical feature abstraction + Multi-source training data
Data format heterogeneity	Unified parsing pipeline with flexible pattern matching
Inconsistent timestamp formats	Canonical conversion to DateTime objects with microsecond precision
Hexadecimal CAN ID variations	Standardized conversion to decimal integers
Defining "normal" behavior	Robust scaling (median/IQR) + 99.5th percentile thresholding
Real-time noise in predictions	Rolling median smoothing (3-5 window size)

4. Experimental Results

The system was evaluated on four distinct attack types representing the spectrum of CAN bus threats:

4.1 Attack Types Evaluated

1. Denial-of-Service (DoS): Flooding attack with high-priority messages
2. Fuzzing Attack: Random ID and payload injection
3. RPM Spoofing: Targeted manipulation of engine RPM signals
4. Gear Injection: Spoofing of gear position messages

4.2 Performance Summary

TABLE I: EXPERIMENTAL RESULTS SUMMARY

Attack Type	Accuracy	Precision	Recall	F1-Score	TP	FP	TN	FN
Fuzzy Attack	97.44%	96.52%	98.80%	97.65%	15,636	563	13,075	190
DoS Attack	97.12%	94.77%	99.75%	97.20%	14,012	773	13,281	35
RPM Spoofing	99.02%	98.00%	99.81%	98.90%	10,680	218	13,414	20
Gear Injection	98.92%	97.80%	99.77%	98.78%	10,639	239	13,404	24

Key Findings:

- Consistent high performance across all attack types (97-99% accuracy)
- Exceptional recall rates (98.8-99.8%) indicating minimal missed attacks
- Strong precision (94.8-98.0%) indicating low false positive rates
- Best performance on RPM Spoofing (99.02% accuracy, 99.81% recall)

4.3 Model Comparison

TABLE II: DETECTION METHOD PERFORMANCE

Detection Method	Avg. Accuracy	Avg. Precision	Avg. Recall	Best For
Autoencoder (AE)	98.63%	97.63%	99.63%	All attacks
One-Class SVM	60.31%	64.82%	99.88%	High recall only
Hybrid (AE $\wedge$ SVM)	98.18%	96.77%	99.53%	Balanced precision/recall

4.4 Feature Analysis Insights

The system's feature analysis revealed clear statistical signatures for different attack types:

- Fuzzy/DoS Attacks: Characterized by drastically reduced message rates (838 vs 1842 msgs/sec) and altered ID distributions
- Spoofing Attacks: Identified by significant changes in dominant ID frequencies (top_id_1_count increased 6x for RPM spoofing)

These findings validate that statistical feature abstraction effectively captures attack signatures regardless of the specific vehicle platform.

5. Validation of Vehicle-Agnostic Capability

The core hypothesis—that statistical feature abstraction enables vehicle-agnostic detection—was validated through:

1. Training on multi-vehicle data: Model learned from four different vehicle platforms
2. Consistent performance: High accuracy maintained across all attack types
3. Generalized patterns: System focused on network behavior rather than specific CAN IDs
4. Cross-platform applicability: Same model worked effectively without vehicle-specific tuning

6. Contributions

This project makes several key contributions to automotive cybersecurity:

6.1 Methodological Contributions

- Vehicle-agnostic detection framework using statistical feature abstraction
- Hybrid anomaly detection combining autoencoders and one-class SVM
- Multi-source dataset processing pipeline for heterogeneous CAN data

6.2 Practical Contributions

- Open-source implementation with complete codebase
- Comprehensive dataset aggregating multiple public sources
- Production-ready architecture with real-time processing capabilities

6.3 Research Contributions

- Demonstrated viability of vehicle-agnostic IDS approach
- Quantified performance across diverse attack types
- Identified key features for effective CAN intrusion detection

7. Limitations and Future Work

While the system demonstrates excellent performance, several areas warrant further investigation:

Current Limitations

1. Payload-agnostic detection: Current features ignore data byte analysis
2. Static thresholding: Fixed anomaly threshold across all driving contexts
3. Single-scale windowing: Fixed 100ms windows may not optimize all attack detection
4. Binary classification: System detects anomalies but doesn't classify attack types

Future Directions

1. Payload-aware features: Incorporate byte entropy and physical consistency checks
2. Adaptive thresholding: Context-aware anomaly detection based on driving state
3. Multi-scale analysis: Parallel processing with different window sizes
4. Attack classification: Supervised classifier for identifying specific attack types
5. Edge deployment optimization: Model compression for resource-constrained ECUs

8. Conclusion

This project successfully demonstrates that vehicle-agnostic intrusion detection is not only feasible but highly effective for automotive CAN networks. By focusing on fundamental statistical properties of network traffic rather than vehicle-specific message semantics, the system achieves:

- High accuracy (97-99%) across diverse attack types
- Excellent recall (98.8-99.8%) with minimal missed attacks
- Low false positive rates (precision 94.8-98.0%)
- True vehicle-agnostic capability without platform-specific tuning

The complete implementation—including all code, datasets, and documentation—is available in the project repository, providing a foundation for further research and practical deployment in next-generation connected and autonomous vehicles.

Repository: https://github.com/Haneen961/vehicle-agnostic-IDS