

# A Comprehensive Survey on Anomaly Detection Techniques in VANETs: Challenges and Opportunities

Manne Naga Chandra Sekhar Chowdhary, Bandaru Rohan Satya Balaji<sup>a</sup>, S Sreenivasa Chakravarthi and S Sountharajan<sup>b</sup>

*Department of Computer Science and Engineering,  
Amrita School of Computing,  
Amrita Vishwa Vidyapeetham, Chennai, India*

**Keywords:** Vehicular Ad Hoc Networks (VANETs), Intelligent Transportation Systems (ITS), Federated Learning, Intrusion Detection Systems (IDS), Data Manipulation Attacks, Denial of Service (DoS), Collaborative Learning, Network Intrusion.

**Abstract:** The emergence and development of Vehicular Ad Hoc Networks (VANETs) as part of Intelligent Transportation Systems (ITS) bring with them critical operational challenges, with security being paramount. Among these, the detection of anomalies stands out as a vital task to ensure the smooth functioning of VANET communication. Anomaly detection, leveraging advanced machine learning (ML) and deep learning (DL) techniques, has emerged as a vital solution to address these challenges. This paper presents a comprehensive survey of recent developments in anomaly detection methods for VANETs. It investigates the supervised, unsupervised, and hybrid learning techniques of CNNs and LSTM networks and federated learning models for anomaly identification in various scenarios. Furthermore, benchmark datasets such as KDD99, NSL-KDD, and VeReMi are reviewed for evaluating the efficacy of these methods. This survey discusses the strengths, weaknesses, and emerging trends within anomaly detection. One such trend is collaborative and privacy-preserving frameworks for anomaly detection. The current work aims to provide guidance for future research in finding robust and real-time anomaly detection systems, thus ensuring the security and reliability of VANETs in environments of increasing complexity.


## 1 INTRODUCTION


VANETs, a key component of Intelligent Transportation Systems, is a transforming technology for the road safety and traffic management era. Self-organizing networks, enabling communication between vehicles (Vehicle-to-Vehicle, V2V) and infrastructure (Vehicle-to-Infrastructure, V2I), have the potential to share real-time information about traffic conditions, possible danger, or other important information. However, despite the promise of VANETs, several critical challenges are found in this area, especially concerning secure and reliable communication. Security issues like Denial of Service (DoS), Sybil, and position falsification attacks can compromise the integrity and functionality of VANET, hence requiring efficient anomaly detection mechanisms to

mitigate these issues.

Anomalies in VANETs manifest as irregular patterns in communication, vehicle behavior, or network infrastructure interactions. Advanced methodologies are required to detect threats in real-time. Application of machine learning and deep learning approaches have been quite promising in this domain. Some of the techniques that are used for detection are Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and ensemble learning frameworks. Benchmark datasets developed for the purpose include KDD99, NSL-KDD, and VeReMi datasets, representing diversity in test and development scenarios.

The evolution of the anomaly detection in VANETs is perpetual, and it has traditionally integrated hybrid frameworks that blend statistical analysis with machine learning models to improve the accuracy of detection while reducing false positives.

<sup>a</sup>  <https://orcid.org/0009-0000-9843-5552>

<sup>b</sup>  <https://orcid.org/0000-0003-4248-3875>

For instance, more recent studies demonstrated collaborative learning techniques in distributed environments can be effective to enable vehicles to collaborate and detect threats. Federated learning and privacy-preserving methods have further contributed to advanced state-of-the-art against data security with robust anomaly detection capabilities.

The literature survey will be aimed at an integrated view of all the recent advances in anomaly detection in VANETs, including strengths, weaknesses, and future research directions. Through the synthesis of findings from highly impacting studies, we contribute further to understanding how emerging technologies can mitigate security threats in these dynamic and decentralized networks.

## 2 RESEARCH METHODOLOGY

The approach to research for this survey paper delves into the recent developments and datasets, considering the fast-growing area of anomaly detection in vehicular ad hoc networks (VANETs). This research seeks to assess the merits and demerits of machine learning and deep learning as well as combined approaches in identifying anomalies, more specifically, misbehaviour in VANETs. There were also metric and non-metric and dataset approaches studied to present the other side of the coin in the scope. Our research process commenced with comprehensive research targeted towards various high-ranking journals and publications including Elsevier, IEEE, Springer and MDPI among other. Such wide-ranging survey of the literature allowed us to synthesize, evaluate and extract important information, thereby converting unrefined data into useful knowledge. A prudent selection of the most appropriate works was made to ensure that the study was adequately grounded.

### 2.1 Research Questions

In this section, we will present the research questions that were the focus of our investigation. These questions were the principal guideline the blueprint of our analysis that helped us not to deviate from the main purpose and issues regarding the anomaly detection in VANETs.

- **RQ1:** *What are the current machine learning-based approaches to misbehaviour detection in VANETs, and what are their drawbacks?*
- **RQ2:** *What remedial measures and approaches can be adopted for data bias mitigation in the con-*

*text of VANETs to improve the effectiveness and equity of detection of different kinds of vehicular misbehaviours?*

- **RQ3:** *How do the malfunctions of physical sensors mounted on vehicles cause the failures of the VANETs?*
- **RQ4:** *What motivates most of the research on misbehaviour in VANETs into position falsification, and why is it of great concern globally to vehicular safety?*
- **RQ5:** *What are the major features or data points that play the greatest role in the effective detection of anomalies in VANETs?*
- **RQ6:** *How can the anomaly detection capability of vehicles in a VANET be improved with the use of collaborative learning techniques?*

These questions help us define the scope of our evaluation of the anomaly detection field as well as provide an overview on its state and trends.

## 3 BACKGROUND ON VANETS

VANETs represent an advanced type of mobile ad-hoc networks (MANETs). They allow vehicles to successfully communicate both with each other as well as with roadside infrastructure. This makes it possible to implement any range of applications: from safety to traffic applications and even infotainment ones. However, the operational definition of VANETs also brings about some odd security issues and threats.

### 3.1 Overview of VANET architecture

VANETs enable vehicle-to-vehicle communication (Vehicle-to-Vehicle or V2V) and vehicle-to-infrastructure communication (Vehicle-to-Infrastructure or V2I) as shown in the figure. Vehicles in VANETs are essentially nodes of a moving network that keeps on changing along with the change in locations. Such nodes exchange important information such as the speed of the vehicle, position, or traffic conditions. Dedicated Short Range Communication protocols, wireless access in vehicular environments, and cellular networks, such as 5G, are the major enablers for VANET communication (Guerrero-Ibáñez et al., 2013).

1. **On-Board Units (OBUs):** On Board Units are installation devices in vehicles, which have capabilities fitted for wireless communications, Global Positioning System, and other sensor systems.

The OBUs make possible Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications, through which certain information like vehicle position, speed, and environmental data can be exchanged (He, 2024).

2. **Roadside Units (RSUs):** RSUs are fixed infrastructure elements placed on roadsides or at intersections. An RSU extends network coverage, allows for communication between distant vehicles, and supplies services such as traffic signal control, toll collection, and access to internet services (Shakir et al., 2024).

### 3.2 Types of communication

- **Vehicle-to-Vehicle (V2V) Communication:** Direct communication among vehicles will enable them to share real-time information about traffic conditions, potential hazards, and other relevant data in order to improve road safety and traffic management.
- **Vehicle-to-Infrastructure (V2I) Communication:** It helps make interaction between vehicles and fixed infrastructure elements such as traffic lights and road signs, as well as central traffic management, possible. Such communication optimizes the flow of traffic, reduces congestion, and increases safety along the roads.

### 3.3 Types of Anomalies

Figure 1 describes a hierarchical classification of anomalies in VANETs according to the nature and impact they have. Anomalies are classified into five major types, including Network-Related Anomalies, Position-Related Anomalies, Speed-Related Anomalies, Data Manipulation, and Communication Anomalies.

Network-related anomalies include DoS attacks, Sybil attacks, Wormhole attacks, Black Hole attacks, and Gray Hole attacks, each designed to violate network integrity through communication. Position-related anomalies include incidents like Constant Position, Random Position, Constant Position Offset, Random Position Offset, and Eventual Stop, all of which hinder the precision with which a vehicle is positioned. Anomalies related to speed include irregularities such as Constant Speed, Random Speed, Constant Speed Offset, and Random Speed Offset, likely creating a problem of miscommunication or even causing a traffic hazard. Data manipulation comprises Data Replay and Disruptive anomalies where the adversary re-sends or manipulates data without authorization to create chaos. Communication

anomalies include Delayed Messages whereby a legitimate message is delayed, which is time sensitive for VANET communications.

It further depicts the multifaceted and wide variety of threats that exist under this hierarchical structure in VANETs, showing emphasis and gravity for strong anomaly detection systems in vehicular networks to ensure security and reliability.

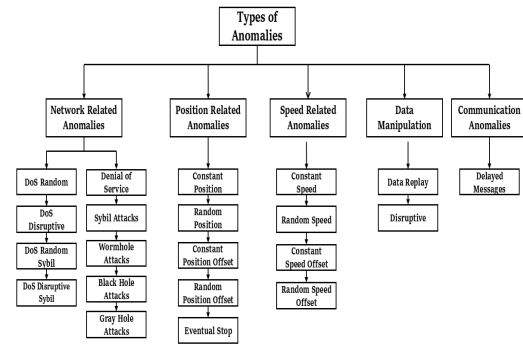


Figure 1: Anomalies Identified in the Survey.

## 4 SURVEY OF RELATED WORK

### 4.1 Network Anomalies Related Work

The paper "Detecting Sybil Attacks Using Proofs of Work and Location in VANETs" (Baza et al., 2022) introduces a new method based on the amalgamation of proofs of work and location to detect Sybil attacks. It uses threshold signatures from RSUs, coupled with vehicle trajectory analysis to prevent RSU compromise attacks. Simulations, based on a dataset motivated by Nashville, TN, have 160 vehicles whereby 10% are malicious entities. A Proof of Work algorithm and maximum clique analysis is used to detect the Sybil nodes. The proposed approach reduces the FNR and detection time up to 50% compared to the previous approaches, while maintaining low communication and computational overhead.

Haowen Tan et al. (Tan et al., 2018) proposed a certificateless authentication scheme integrated with unsupervised anomaly detection for VANETs, which focuses on DoS attacks and traffic flow anomalies. The scheme makes use of certificateless cryptography with no pairing operations. It uses the Chinese Remainder Theorem for efficient group key distribution and dynamic time warping to detect anomalies. Security analysis and evaluation of performance attest that the approach results in strong security as well as efficiency better than the existing ones.

Nikita Lyamin et al. (Lyamin et al., 2018) presented real-time detection of jamming DoS attacks in VANET, with a focus on platooning application. A hybrid detection approach coupling the statistical analysis with data mining techniques is proposed to identify jamming in real time even with random jitter in cooperative awareness messages. The simulation results have shown that the method is efficient for detecting both random and ON-OFF jamming strategies over different sizes of the platoon.

Nie et al. (Nie et al., 2019) proposed a generalized anomaly detection framework for VANETs targeting PHY-layer spoofing, jamming, and DDoS attacks. The method exploits spatiotemporal traffic characteristics and sparsity modeled using a CNN architecture. A loss function based on Mahalanobis distance and reinforcement learning enhances the precision of detection. Although the specifics of the dataset are not mentioned, the simulations validate the method. The framework achieves over 90% accuracy and precision across various observation durations, with low false alarm rates, demonstrating robust and reliable anomaly detection.

Shu et al. (Shu et al., 2021) performed an experiment that utilized CIDS on two datasets: KDD99 and NSL-KDD. From the output, it is easy to infer that collaborative CIDS perfectly outperforms IndiDetection without fairly having any higher system overheads while nearly performing equally well as Cent-Detection. In the case of the KDD99 dataset, CIDS attains a surprisingly high accuracy level, which is 98.38% for SDN3, precision showing efficient detection capabilities within varied network scenarios (95.14% for SDN3). For the NSL-KDD dataset, CIDS is robust in all aspects with an accuracy reaching 96.75% for SDN3, precision at 91.83%. These metrics imply that CIDS is efficient in terms of achieving balance between detection accuracy and complexity with the way it reduces computational and communication, hence making it feasible for this kind of distributed intrusion detection within VANETs.

The paper, "Intelligent Hierarchical Security Framework for Vehicular Ad Hoc Networks"\* (Goncalves et al., 2021) introduces a multi-level architecture to enhance VANET security and attack detection. The framework consists of four levels: individual vehicles (L0), vehicle clusters (L1), roadside units (L2), and backend servers (L3). Each layer leverages localized machine learning; for instance, L0 relies on lightweight decision stumps, L2 adopts Random Forest to achieve better results, and L3 resorts to the ensemble technique, such as MLP or J48, but with higher complexity. The VPKIbrID model ensures safe and confidential communication between

entities. Custom datasets, developed using SUMO and NS-3, have shown effective results in detecting DoS attacks. This hierarchical approach integrates security and attack detection efficiently, while each node takes benefit from its strong capabilities.

The paper, "A New Multivariate Approach for Real Time Detection of Routing Security Attacks in VANETs (Ajjaj et al., 2022) proposes a Multivariate Statistical Detection Scheme (MVSDS) for detecting black hole attacks in VANETs. The method applies multivariate normality tests without changing the existing routing protocols, therefore considering the traffic metrics like throughput, dropped packet ratio, and overhead traffic ratio. Techniques used are Min-Max Normalization and tests of the kind of Rao-Ali and Ryan-Joiner for anomaly detection. Simulations using SUMO and NS-3 confirm that MVSDS effectively identifies black hole attacks as it is an efficient technique with a high sensitivity and which detects performance degradation, namely, reduced throughput and increased packet loss.

Abderrahim Benslimane et al. (Agrawal et al., 2022) proposed a deep learning-based intrusion detection system (IDS), NovelADS, that was designed to detect anomalies in intra-vehicular networks using the CAN protocol. Its focus attacks include DoS, Fuzzy attacks, RPM Spoofing, and Gear Spoofing. NovelADS uses sequence-level classification and spatio-temporal analysis of legitimate network messages to detect deviations causing attacks. NovelADS uses techniques such as thresholding and error reconstruction along with a novel statistical method for automated threshold determination, which reduces manual intervention. Different architectures of neural networks are trained to enhance detection efficiency, especially in the case of Fuzzy attacks, exhibiting near-perfect precision (0.9995), recall (0.9991), and F1-score (0.9993). This approach performs well in cases of other attack types as well by outperforming traditional approaches.

Among the attacks identified in the paper (Shams and Ulusoy, 2020) on DoS attacks in VANETs is the intrusion detection system using a Support Vector Machine that consists of packet dropping and delaying attacks which disturb the communication in the network and bring risks to vehicular safety. This experiment utilizes a custom dataset created using computer simulation of both normal and malicious patterns of traffic in a realistic mobility vehicular environment. The proposed IDS, in fact, uses SVM as a core algorithm that performs anomaly detection. The detection of probable intrusions takes place at the receiving vehicle by analyzing the packet arrival pattern, and, in this case, feature selection is conducted



to identify those attributes which are critical for DoS attacks detection. Results from SVM-based approach are compared with other classifiers; the method used here signifies major improvement.

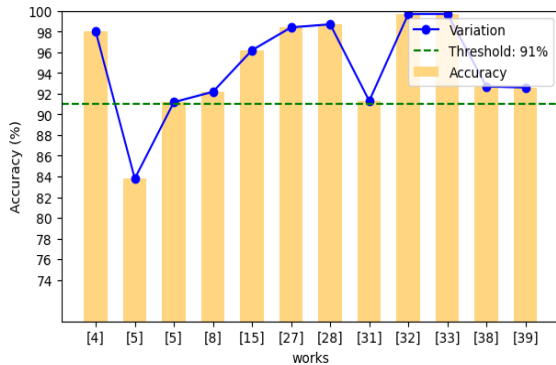


Figure 2: Accuracy achieved in various works

The paper by Nabil Nissar et.al(Nissar et al., 2024) It is stated that there exists a wide range of cyber threats in VANETs, specially zero-day attacks and dynamic anomalies. A framework of Variational Autoencoders is presented for anomaly detection, and the objective functions dealt within this present work comprise of KL-divergence together with reconstruction error, and two optimization algorithms are used- AGE-MOEA and R-NSGA-III. Due to the carried study, it has been mentioned that the accuracy and precision rate for the R-NSGA-III has achieved 90.02% and 92.98% respectively. The results in case of AGE-MOEA are also competitive enough with an accuracy of 88.37% and a precision rate of 92.69%.

In order to hinder wormhole and black hole attacks in VANET environments, Boya Liu et al.(Liu et al., 2023) presented federated learning that incorporates a reward mechanism with trust values. Apart from this, the approach utilizes homomorphic encryption for defense against privacy breaches and attacks. The research work has allocated lossless tree enhancement technique that achieves its goal via federated learning, which has put emphasis on methods to select nodes and model gradients aggregation.

Gurtej Kaur et al. (Kaur et al., 2022) presents another work on the analysis of the AODV routing protocol to show the performance of Gray hole attacks in VANETs by considering some metrics such as throughput, PDR, NRL, and delay. The algorithm developed is for Nack with Smart Neighbourhood-Hole Recovery (NSN-H) with respect to the ensured message transmission with efficient reliable communication while considering the minimal possible delay.

Thuvva Anjali et al. (Anjali et al., 2024) pro-

posed a two-tier strategy that mitigates DoS attacks in VANET by targeting two types of threats: external and internal threats. This strategy enhances security by allowing the verification of identities of communication entities through the creation of signatures derived from private and public keys.

M Poongodi et al. (Poongodi et al., 2019) proposed a trust-based framework to mitigate the issues confronted in VANETs due to DDoS attacks, wherein it clusters the nodes with trust scores and a genetic algorithm can be put in place to adaptively structure in a hierarchical fashion to better improve routing efficiency in the light of identifying malicious nodes. It has significant network performance improvements.

The KDD Cup 99, NSL-KDD, CICIDS2017, and UNSW-NB15 benchmark datasets have been used in many studies for developing and evaluating models for intrusion detection, anomaly detection, and misbehaviour detection in VANET communications with the purpose of mitigating network-related threats. Because of the heavy use of a huge dataset, these datasets contain exhaustive ranges of attack scenarios and normal traffic patterns.

The work by Alsarhan et al. (Alsarhan et al., 2021) presents an SVM-based intrusion detection system for VANETs optimized using Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO). It ensures the experimental results over the KDD99 dataset with the accuracy of 98% and the detection rate of 99%. Such optimizations enhance the accuracy of the SVM classifier. KDD99 dataset is the widely used benchmark to assess intrusion detection systems. It includes both regular activities and different types of attacks such as DoS, probing, User to Root (U2R), and Remote to Local (R2L) attacks in synthetic data of the network traffic.

An anomaly-based system for VANETs by using K-means clustering and fuzzy set theory was presented in a study by Rafsanjani et al. (Kuchaki Rafsanjani et al., 2021). It includes a 0.97 detection rate for DoS attacks, with 0.99 precision and an F-measure value of 0.98. The integration of the two modules of clustering and fuzzy logic supports effective anomaly identification.

In research study by AlMahadin et al. (Aoudni et al., 2024) talks about anomaly detection in VANET network traffic using a GRU-based deep learning model in evaluating the SEMI-GRU with the NSL-KDD dataset. The outcome of the experiment reveals that the 5-layer and 8-layer models are performing the best, with the highest accuracy being provided by the 5-layer model as 83.79%. SMOTE oversampling technique has been utilized for tackling class imbal-

ance within the dataset. However, the method presented in this work is computationally intensive.

In Gyawali et al. (Gyawali et al., 2020), a method that used a fusion of Random Forest and Dempster-Shafer theory for misbehaviour detection in vehicular networks is proposed. The labelled datasets were obtained by simulating the Veins framework in OMNET++. It achieved a precision of 0.99, recall of 0.96, and an F1-score of 0.97, particularly in constant position attacks. On the other hand, the conclusion of the study is that although the cryptographic method seems quite effective, due to the susceptibility of vehicular networks to internal attacks, even the legitimate network nodes can become a threat to the overall network. The detection also seems pretty dependent upon feedback received from the vehicles.

Bangui et al. (Bangui et al., 2021) proposes a hybrid data-driven model for intrusion detection in VANETs. It combines two important parts: firstly, making use of a classification algorithm, identifies known attacks, and secondly, utilises an anomaly detection approach based on the coresets technique to filter dishonest nodes from being considered within cluster-based filtering. This two-phased strategy reflected an excellent accuracy of 96.93% and F1-score of 94.41%, marking a tremendous progress for the real-time IDS of VANETs. However, it was not that comprehensive to include all kinds of attacks in VANETs and therefore would require greater refinements to fight more complex attack scenarios.

Baharlouei et al. (Baharlouei et al., 2024) designed a real-time anomaly and attack detection system in the VANET using XGBoost and federated learning. Each vehicle trains an XGBoost model locally, enabling efficient detection of malicious behavior and attacker identification. It provided a detection rate of 99.66% along with a false negative rate of 0.72%. Simulations were performed for several cities with attacker densities ranging from 5% to 30% in each; the simulations ran for 3600 seconds, which is equivalent to 24 hours. Although this method has promising results, applying it to real-world applications is quite complex due to the variety of vehicular environments.

Kumar et al. (Kumar and Chilamkurti, 2014) introduced T-CLAIDS, a Trust-based Collaborative Intelligent Intrusion Detection System for detecting malicious activities in VANETs. The system uses a Collaborative Trust Index (CTI) to enhance its classifier's effectiveness across various attack scenarios. It combines Learning Automata (LA) for vehicle state monitoring and a Markov Chain Model (MCM) to model state transitions, achieving a high detection rate of 99% and a packet delivery ratio of 98%. While ef-

fective in dynamic vehicular environments, its performance may degrade in highly variable or sparse network conditions and against sophisticated attack types.

Garg et al. (Garg et al., 2019) designed Sec-IoV, a multi-stage anomaly detection scheme in the Internet of Vehicles (IoV). The authors have employed a hybrid optimization of an SVM classifier's parameters to boost the accuracy of the classifier with an operator derived from the mutation of the Artificial Bee Colony optimization, namely, Cauchy-based mutation operator (C-ABC). Simulations were performed over OMNET++ and SUMO environments with respectively high detection rates and accuracy. However, Sec-IoV faces scalability challenges and has been validated only in limited real-world scenarios, raising concerns about its effectiveness against diverse attack types.

Table 1: Overview of Intrusion Detection Systems in VANETs

Work	Method	Results	Strengths	Weaknesses
[3]	Semi-supervised VAEs for intrusion detection	99.46% Accuracy	Robust to attack strategies	Struggles with novel attacks
[4]	SVM optimized with GA, PSO, ACO	98% Accuracy	Enhanced precision	Limited applicability to real-world data
[5]	GRU with SMOTE for class imbalance	83.79% Accuracy	Effective for sequential data	High computational demand
[6]	Random Forest with Dempster-Shafer theory	0.99 Precision	Effective for specific attack detection	Reliance on feedback data
[9]	Federated learning with XGBoost	99.66% Accuracy	Data privacy, distributed detection	Complexity in diverse environments
[15]	Hybrid data-driven model	96.93% Accuracy	Reduces false positives	May not cover all attack types
[23]	Statistical analysis for jamming detection	Efficient in jamming detection	Robust to random jitter	Limited to simple attack scenarios
[25]	Spatiotemporal modeling with CNN	90% Accuracy	Integrates reinforcement learning	Lacks dataset details
[27]	CIDS evaluated on KDD99	98.38% Accuracy	Effective for distributed detection	Degrades in complex scenarios
[29]	Proof of Work and Location for Sybil attacks	Reduced FNR by 50%	Effective against RSU compromises	Requires reliable RSU network
[36]	NovelADS for fuzzy attacks detection	Precision 0.9995	Near-perfect detection rates	Needs adaptation for real-world use

## 4.2 Position anomalies Related work

Yang et al.(Yang et al., 2023) discusses Connected and Autonomous Vehicles GPS spoofing detection by using Learning from Demonstration (LfD) framework along with Maximum Entropy Inverse Reinforcement Learning (ME-IRL) to model normal driving behavior with decision tree classifier based on objective ratio and trajectory displacement as feature for identifying anomalous behavior. The approach is tested on both KAIST and Michigan datasets where strong performance is seen in terms of detection with low false positive and false negative rates and prove to be robust against stealthy attacks. However, it can detect known attacks and relies on labeled attack data for training and focuses on detection but does not have any mitigation strategies.

Secil Ercan et al. (Ercan et al., 2022) present a distributed intrusion detection system for position falsification attacks in VANETs, such as Constant, Constant Offset, Random, Random Offset, and Eventual Stop attacks. The features presented in the system are AoA, RSSI-based estimated distance, and declared vs. estimated distances, which advance the system with much better detection of events. Using the VeReMi dataset the improved accuracy, F1-score, and computation time across different types of attacks and traffic densities. The machine learning techniques kNN and Random Forest in a Stacking ensemble yield an accuracy of 83.6% in low, 91.5% in medium, and 92.2% in high traffic density conditions.

In the paper "Anomaly Detection for Internet of Vehicles: A Trust Management Scheme with Affinity Propagation", Shu Yang et.al.(Yang et al., 2016) proposed an anomaly detection scheme in IoV. The two types of anomalies that are targeted by the study are, on one hand, malicious vehicles the intent of which is harmful and on the other incapable vehicles whose action causes disorder without their intent. The Cluster-Based Anomaly Detection introduces cluster-based and central reputation components for managing trust dynamically and in the long run. The CAD approach generally involves cluster building, abnormal behavior detection within clusters, and election of cluster heads for managing trust management. It uses affinity propagation as the foundation for clustering along with trust evaluation. Simulation proved that this system had a very low failure rate in detecting abnormal vehicles with less than 1%.

Alladi et al. (Alladi et al., 2021) presented a deep learning-based framework, DeepADV for anomaly detection in VANETs. The CNN-LSTM configuration was found to be highly accurate with 98.4% accuracy in faults, 98.7% in attacks, and 98% in com-

bined anomalies. DeepADV uses sequence reconstruction via a thresholding algorithm along with features like position coordinates and speed. The framework successfully detects subtle anomalies such as Constant Position Offset faults, Delayed Messages, and Eventual Stop attacks. This performance was validated through extensive experimentation over different types of anomalies, signifying the adaptability to unknown anomalies without the need for retraining. Furthermore, the deployment on Nvidia Jetson Nano hardware further illustrates its real-time applicability in VANET environments.

Steven So et al. (So et al., 2018) presents a framework in which plausibility checks are incorporated as a feature vector for the machine learning models, SVM and KNN, to enhance the accuracy of misbehavior detection by as much as more than 20% within a recall of within 5%. They introduced new features like average distance and average velocity, which are derived from the distance reported by the GPS location and reported velocities by the sender.

What Pranav Kumar Singh et al.(Singh et al., 2019) added to the features were the difference between the sender and receiver positions/speed in detecting certain types of attacks. The authors found out that position was important for differentiating attackers from legitimate vehicles, since all types of attacks were based on position falsification, which caused different trends in the values of position

The Hybrid Position Forger Attack Detection algorithm by Shahid (Shahid and Jaekel, 2023) introduces a novel hybrid approach to detecting position forgery attacks in Connected Vehicles, making use of the Veremi dataset. By combining the techniques of machine learning with plausibility checks, the HP-FAD approach increases the accuracy of detection compared to earlier existing approaches and yields a considerably high F1-score of 99.40%. This approach precisely targets attacks that come in the form of position forgery attacks in BSMs; thus, it is an effective solution to targeted anomaly detection.

Other authors (Behravan et al., 2022) used stacking ensemble learning to improve misbehaviour detection in VANETs by using the Veremi dataset. In this research, five kinds of position falsification attacks: constant, constant offset, random, random offset, and eventual stop attacks-have been addressed, which ultimately gives a holistic view of the threat. In fact, the authors come up with two stacking-based detection systems: one is based on traditional classifiers (Logistic Regression, K-NN, Decision Trees, Naive Bayes) and the other on neural networks. Each of the above is designed to classify misbehaving vehicles with higher accuracy by considering the fact that

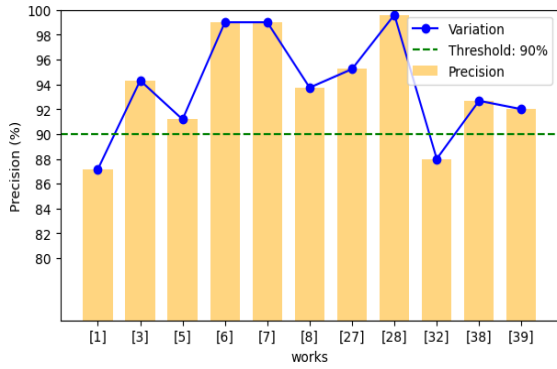


Figure 3: Precision values achieved in various works.

each attack is different.

In this study, Liu (Liu, 2022) has investigated the application of LSTM networks toward detecting misbehaviour in VANETs. The study utilizes the Veremi dataset and would focus on developing classification models to identify different types of communication anomalies using deep learning models. This is quite a different approach from traditional machine learning as proposed thus far to identify and characterize different types of improper communication behaviors in vehicular networks.

A new approach of classification called the One vs. All Binary Tree (OVA-BT) was presented by Slama (Slama et al., 2023) to deal with imbalanced datasets in detecting misbehaviour in VANETs with accuracy metrics across the Veremi dataset. In this method, the approach uses a binary classifier to any class type of misbehaviour and during the classification phase, ensures both classes are dealing with equal importance, majority and minority classes. The study has indicated remarkable improvements in accuracy in most classifiers by using the OVA-BT approach, as KNN\_OVA-BT, SVM\_OVA-BT, and RF\_OVA-BT have improved precision values up to 10%, 20%, and 4%, respectively, compared with traditional approaches.

Bayan's (Bayan et al., 2024) work talks about a decentralized Deep Learning-based Intrusion Detection System (DL-IDS). This system utilizes a Multi-Layer Perceptron MLP for detecting the position falsification attacks occurring in inter-vehicle networks, using the dataset of Veremi. The proposed system successfully detects multiple position falsification attacks and achieves F1 scores of 93, 94, and 92 for different attack scenarios. The innovative features at its core are the aggregation of RSSI from first-hop neighbors and TDoA, which are found to be essential in improving false position detection accuracy and strengthening the security of VANETs.

Table 2: Research Highlights on Detection Techniques for Position Falsification Attacks

Work	Method	Strengths	Weaknesses
[34]	Distributed intrusion detection for position falsification attacks	Novel features improve detection capability	Limited to specific attack types
[28]	Cluster-Based Anomaly Detection in IoV	Dynamic trust management	Complexity in real-world scenarios
[28]	DeepADV framework with DNN architectures	High adaptability for various anomalies	Some anomalies are subtle and may go undetected
[1]	SVM and KNN with plausibility checks	Effective feature enhancement	Limited to specific vehicle scenarios
[2]	Position/speed difference as a feature	Improved detection capability	Potential for false positives
[42]	Hybrid Position Forger Attack Detection (HPFAD)	Robust against position forgery attacks	Targeted approach limits generalizability
[43]	Stacking ensemble learning for misbehavior detection	Comprehensive solution for security threats	Complexity in classifier selection
[44]	LSTM networks for misbehavior detection	Advanced categorization of behaviors	Requires substantial data for training
[46]	One vs. All Binary Tree (OVA-BT) method	Effective for imbalanced datasets	Performance variability across classifiers
[47]	Deep Learning-based Intrusion Detection System (DL-IDS)	Effective use of innovative features	May struggle with new attack vectors

### 4.3 Hybrid

Zaidi et al. proposed in (Zaidi et al., 2016) a host-based IDS specifically for detecting nodes that produce false information attacks in VANETs. This system processes the received data based on statistical methods instead of trust or reputation metrics. Simulations are conducted using OMNET++, SUMO, and VACaMobil, with various traffic and rogue node configurations. Key parameters of vehicle such as speed and density are analyzed. The IDS utilizes hypothesis testing, utilizing the t-test to compare the received parameters against expected values based upon Greenshield's traffic flow model. Thus, data collection can be cooperative among vehicles, and a distributed IDS architecture can be realized. Simulation results demonstrate this technique's effectiveness, including a high true positive rate in the presence of up to 30% rogue nodes and a false positive rate that remains low at up to 20% of rogue nodes, surpassing existing techniques.

Omessaad Slama et al. (Slama et al., 2022) in their study discussed feature selection methods, such as Recursive Feature Elimination, F-test Anova to avoid



overfitting and improve the model’s generalisability. The authors proposed a Guided Learning Approach for Multi-class Classification (G-LAMC) for addressing class imbalance issues. The author pointed that the Random Forest algorithm obtained better results than any of the models used in the study.

Chen-Khong Tham et al. (Tham et al., 2023) performs a research towards the application of federated learning techniques, FedAvg-SGD, FedAvg-Adam and FedProx toward anomaly detection in vehicular networks utilizing the VeReMi Extension dataset. The research shows efficacy with federated learning, achieving accuracy of up to 92.18%, while FedAvg-Adam showed precision at 93.74%, recall at 92.43%, and an F1-score of 93.08%. IID and non-IID distributions on federated learning models: To understand this, the authors simulate different data distributions - some randomly and others based on quadrants. However, the study is somewhat limited in that it only considers a binary classification scenario and distinguishes between normal and anomaly classes.

Devika S et al. (S et al., 2024) has proposed a new unsupervised anomaly detection framework known as VADGAN. It works in collaboration with combined GANs and LSTM to facilitate effective anomaly detection in CAVs. Using the Veremi extension dataset, it achieved recall at 81.838%, thus establishing one among many other attack types in this kind of scenario. The study compares different architectures, such as LSTM, RNN, and GRU, and it has been observed that the outcome of LSTM was excellent in order to detect complex anomalies. This approach is very timely as it represents an enhanced development with the use of GAN-based models for the improvement of CAVs’ security.

5 SURVEY WORK

Figure 2 illustrates the accuracy percentages achieved by the different models during anomaly detection; in other words, the comparison of how each of these models may be used to determine the effectiveness of anomalies. Figure 3 represents the precision percentages obtained by the models, which denotes the true positive instances of anomalies that are detected by them. Precision is crucial, as it is a measure of how reliable the models are, the ratio of relevant instances retrieved.

Table 3: Summary of Recent Advances in Misbehaviour and Malfunction Detection Works

Work	Method	Strengths	Weaknesses
[30]	Host-based IDS for rogue node detection in VANETs	Higher true positive rate; effective anomaly detection	Limited to specific rogue node scenarios
[3]	Guided Learning Approach for Multi-class Classification	Improved model generalizability	Potential overfitting with small datasets
[8]	Federated learning techniques for anomaly detection	High accuracy and precision; insights into data distributions	Limited to binary classification scenarios
[10]	VADGAN: GANs with LSTM for anomaly detection	Effective in identifying various attack types	Limited to specific architectures in performance comparison

6 SUMMARY OF REVIEW FINDINGS

**Response to RQ1.** Current machine learning-based methodologies demonstrate considerable effectiveness in detecting misbehavior within VANETs. Techniques such as Support Vector Machines (SVMs), deep learning frameworks (e.g., Long Short-Term Memory networks), and ensemble learning approaches have shown promising results. However, key limitations exist, primarily due to reliance on inadequate datasets that are far away from the real-world scenarios, leading to restricted model adaptability. Furthermore, challenges such as computational complexity, dependence on predefined attack databases, and the generalizability of results across various datasets considerably impede overall performance.

**Response to RQ2.** To rectify class imbalances in data and ensure more accurate detection of vehicular misbehavior, several strategic interventions can be implemented:

- **Resampling Techniques:** Utilizing SMOTE (Synthetic Minority Over-sampling Technique), DSSTE and ADASYN (Adaptive Synthetic Sampling) for oversampling minority classes, while employing random and cluster-based under-sampling.
- **Advanced Algorithms:** Implementing One-vs-All Binary Tree classifiers (OVA-BT) and utilizing ensemble learning techniques can enhance detection performance across imbalanced datasets.

- **Cost-Sensitive Learning:** Modifying algorithms to include penalties for misclassifications in minority classes as well as weighted loss functions that emphasize minority class errors can help balance detection capabilities.

**Response to RQ3.** The failure of physical sensors in vehicles significantly contributes to the generation of false data in the VANETs. Malfunctions due to wear, environmental factors, or electrical issues lead to erroneous readings, thus compromising the integrity of transmitted data. These sensors are responsible for monitoring and transmitting vital information such as position, speed. When a sensor fails, the data it produces can become inaccurate or completely erroneous. For instance, a malfunctioning GPS sensor might produce incorrect longitude and latitude coordinates, leading to position malfunctions. Similarly, speed malfunctions can occur. Overall, the failure of physical sensors in vehicles introduces significant vulnerabilities in VANETs, as the integrity of the data these networks rely on is compromised. So, it is very important to detect these vulnerabilities in the network.

**Response to RQ4.** Position falsification happens to be a primary focus for researchers, in detection of misbehaviour in VANETS, due to its potential adverse effects on safety and system trust. Accurate positioning is obviously crucial for navigation and collision avoidance, when an attacker manipulates positional data, the integrity of these safety features are compromised. The implications of effectively detecting position falsification extend globally, influencing enhancements in vehicular safety, integration with intelligent/smart city infrastructures, and compliance with SAE regulations. Improved detection methods foster public confidence in vehicular communications and increase the usage of autonomous vehicles.

**Response to RQ5.** Key features or data points that contribute most significantly to accurate anomaly detection in VANETs include:

1. **Position and Movement Data:** GPS coordinates, Vehicle speed, Acceleration, Direction of travel, Difference between sender and receiver positions/speed.
2. **Temporal Features:** Timestamp of messages, Frequency of message transmission, Time intervals between consecutive messages.
3. **Network-related Features:** Signal strength, Packet delivery ratio, Network traffic patterns,

Communication range.

**Response to RQ6.** The utilization of collaborative learning approaches, especially federated learning, offers substantial benefits for enhancing anomaly detection across multiple vehicles within VANETs. By enabling decentralized training, federated learning allows vehicles to share knowledge while keeping sensitive data localized, thus preserving privacy. This method enhances anomaly detection by leveraging the diversity of data collected across various vehicles. The scalability of federated learning means that as more vehicles participate, the model's adaptability to various driving conditions improves. However, challenges such as data heterogeneity and communication overhead must be managed to realize its full potential in detecting anomalies.

## 7 CONCLUSION

This paper gives a detailed analysis of anomaly detection in VANETs; that is, the techniques used to detect anomalies in VANETs. Thus, their importance in ensuring secure and reliable communication within ITS can be emphasized. Different machine learning, deep learning, and hybrid methods analyzed their strengths and weaknesses in identifying misbehavior and anomalies in VANETs. Including benchmark datasets, such as KDD99, NSL-KDD, and VeReMi, showed considerable improvement over the development of accurate detection models, yet the aforementioned drawbacks remain.

Emerging trends are also seen in the application of federated learning and privacy-preserving frameworks, which lead to potential improvement in anomaly detection capabilities and align with data privacy concerns. Class imbalances can be addressed by advanced resampling techniques, while federated learning is part of collaborative learning approaches that further strengthen detection frameworks.

Future directions should include increasing real-time usability, scalable systems in a wide variety of vehicular scenarios, and enriching detection mechanisms for sophisticated attack types. By embracing these new possibilities, VANETs can achieve robust and secure measures for safer and more efficient vehicular networks in increasingly complicated environments.

## REFERENCES

- Agrawal, K., Alladi, T., Agrawal, A., Chamola, V., and Benslimane, A. (2022). Novelads: A novel anomaly

- detection system for intra-vehicular networks. *IEEE Transactions on Intelligent Transportation Systems*, 23(11):22596–22606.
- Ajjaj, S., el Houssaini, S., Mustapha, H., and Houssaini, M.-A. (2022). A new multivariate approach for real time detection of routing security attacks in vanets. *Information*, 13:282.
- Alladi, T., Gera, B., Agrawal, A., Chamola, V., and Yu, F. R. (2021). Deepadv: A deep neural network framework for anomaly detection in vanets. *IEEE Transactions on Vehicular Technology*, 70(11):12013–12023.
- Alsarhan, A., Alauthman, M., Alshdaifat, E., Al-Ghuwairi, A.-R., and Al-Dubai, A. (2021). Machine learning-driven optimization for svm-based intrusion detection system in vehicular ad hoc networks. *Journal of Ambient Intelligence and Humanized Computing*, 14.
- Anjali, T., Goyal, R., and G.N, B. (2024). Prevention of attacks in vehicular adhoc networks. In *2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, pages 1–8.
- Aoudni, Y., Shabaz, D. M., Agrawal, A., Yasmin, G., Alomari, E. S., Al-Khafaji, H. M. R., Dansana, D., and Maaliw III, R. (2024). Vanet network traffic anomaly detection using gru-based deep learning model. *IEEE Transactions on Consumer Electronics*, 70:4548–4555.
- Baharlouei, H., Makanju, A., and Zincir-Heywood, N. (2024). Advent: Attack/anomaly detection in vanets.
- Bangui, H., Ge, M., and Buhnova, B. (2021). A hybrid data-driven model for intrusion detection in vanet. *Procedia Computer Science*, 184:516–523. The 12th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 4th International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops.
- Bayan, S., Mohammad, U., and Al Mohammad, A. (2024). Position falsification attack detection in inter-vehicle networks using deep learning. pages 621–626.
- Baza, M., Nabil, M., Mahmoud, M. M. E. A., Bewermeier, N., Fidan, K., Alasmay, W., and Abdallah, M. (2022). Detecting sybil attacks using proofs of work and location in vanets. *IEEE Transactions on Dependable and Secure Computing*, 19(1):39–53.
- Behravan, M., Zhang, N., Jaekel, A., and Kneppers, M. (2022). Intrusion detection systems based on stacking ensemble learning in vanet. In *2022 5th International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*, pages 1–7.
- Ercan, S., Ayaida, M., and Messai, N. (2022). Misbehavior detection for position falsification attacks in vanets using machine learning. *IEEE Access*, 10:1893–1904.
- Garg, S., Kaur, K., Kaddoum, G., Gagnon, F., Kumar, N., and Han, Z. (2019). Sec-iov: A multi-stage anomaly detection scheme for internet of vehicles. pages 37–42.
- Goncalves, F., Macedo, J., and Santos, A. (2021). An intelligent hierarchical security framework for vanets. *Information*, 12:455.
- Guerrero-Ibáñez, J. A., Flores-Cortés, C., and Zeadally, S. (2013). *Vehicular Ad-hoc Networks (VANETs): Architecture, Protocols and Applications*, pages 49–70. Springer London, London.
- Gyawali, S., Qian, Y., and Hu, R. Q. (2020). Machine learning and reputation based misbehavior detection in vehicular communication networks. *IEEE Transactions on Vehicular Technology*, 69(8):8871–8885.
- He, A. (2024). Understanding on-board units (obu) in vehicle telematics. *Medium*.
- Kaur, G., Khurana, M., and Kaur, A. (2022). Gray hole attack detection and prevention system in vehicular adhoc network (vanet). In *2022 3rd International Conference on Computing, Analytics and Networks (ICAN)*, pages 1–6.
- Kuchaki Rafsanjani, M., Fatemidokht, H., Balas, V. E., and Batth, R. S. (2021). An anomaly detection system based on clustering and fuzzy set theory in vanets. In Balas, V. E., Jain, L. C., Balas, M. M., and Shahbazova, S. N., editors, *Soft Computing Applications*, pages 399–407, Cham. Springer International Publishing.
- Kumar, N. and Chilamkurti, N. (2014). Collaborative trust aware intelligent intrusion detection in vanets. *Computers & Electrical Engineering*, 40.
- Liu, B., Liu, X., Gao, S., Yu, B., and Zuo, P. (2023). Federated learning for vanet based on homomorphic encryption. In *2023 Cross Strait Radio Science and Wireless Technology Conference (CSRSWTC)*, pages 1–3. IEEE.
- Liu, X. (2022). Misbehavior detection based on deep learning for vanets. In *2022 International Conference on Networks, Communications and Information Technology (CNCIT)*, pages 122–128.
- Lyamin, N., Kleyko, D., Delooz, Q., and Vinel, A. (2018). Ai-based malicious network traffic detection in vanets. *IEEE Network*, 32(6):15–21.
- Nie, L., Wu, Y., Wang, H., and li, y. (2019). Anomaly detection based on spatio-temporal and sparse features of network traffic in vanets. *IEEE Access*, 7:177954–177964.
- Nissar, N., Naja, N., and Jamali, A. (2024). Securing vanets: Multi-objective intrusion detection with variational autoencoders. *IEEE Transactions on Consumer Electronics*, 70(1):3867–3874.
- Poongodi, M., Hamdi, M., Sharma, A., Ma, M., and Singh, P. K. (2019). Ddos detection mechanism using trust-based evaluation system in vanet. *IEEE Access*, 7:183532–183544.
- S, D., Shrivastava, R. R., Narang, P., Alladi, T., and Yu, F. R. (2024). Vadgan: An unsupervised gan framework for enhanced anomaly detection in connected and autonomous vehicles. *IEEE Transactions on Vehicular Technology*, 73(9):12458–12467.
- Shahid, M. A. and Jaekel, A. (2023). Hybrid approach to detect position forgery attacks in connected vehicles. In *2023 14th International Conference on Network of the Future (NoF)*, pages 47–51.
- Shakir, A., Islam, M., Mandeep, J., Islam, M., Abdullah, N., Taher, Y., Abdullahi, O., and Soliman, M. (2024). Systematic review of data exchange for road side unit in a vehicular ad hoc network: coherent

- taxonomy, prominent features, datasets, metrics, performance measures, motivation, opportunities, challenges and methodological aspects. *Discover Applied Sciences*, 6.
- Shams, E. and Ulusoy, A. (2020). Performance analysis and comparison of anomaly-based intrusion detection in vehicular ad hoc networks. *Radioengineering*, 29:664–671.
- Shu, J., Zhou, L., Zhang, W., Du, X., and Guizani, M. (2021). Collaborative intrusion detection for vanets: A deep learning-based distributed sdn approach. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):4519–4530.
- Singh, P. K., Gupta, S., Vashistha, R., Nandi, S. K., and Nandi, S. (2019). Machine learning based approach to detect position falsification attack in vanets. In Nandi, S., Jinwala, D., Singh, V., Laxmi, V., Gaur, M. S., and Faruki, P., editors, *Security and Privacy*, pages 166–178, Singapore. Springer Singapore.
- Slama, O., Alaya, B., and Zidi, S. (2022). Towards misbehavior intelligent detection using guided machine learning in vehicular ad-hoc networks (vanet). *Inteligencia Artificial*, 25:138–154.
- Slama, O., Tarhouni, M., Zidi, S., and Alaya, B. (2023). One versus all binary tree method to classify misbehaviors in imbalanced veremi dataset. *IEEE Access*, 11:135944–135958.
- So, S., Sharma, P., and Petit, J. (2018). Integrating plausibility checks and machine learning for misbehavior detection in vanet. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 564–571.
- Tan, H., Gui, Z., and Chung, I. (2018). A secure and efficient certificateless authentication scheme with unsupervised anomaly detection in vanets. *IEEE Access*, 6:74260–74276.
- Tham, C.-K., Yang, L., Khanna, A., and Gera, B. (2023). Federated learning for anomaly detection in vehicular networks. In *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, pages 1–6.
- Yang, S., Liu, Z., Li, J., Wang, S., and Yang, F. (2016). Anomaly detection for internet of vehicles: A trust management scheme with affinity propagation. *Mobile Information Systems*, 2016:1–10.
- Yang, Z., Ying, J., Shen, J., Feng, Y., Chen, Q. A., Mao, Z. M., and Liu, H. X. (2023). Anomaly detection against gps spoofing attacks on connected and autonomous vehicles using learning from demonstration. *IEEE Transactions on Intelligent Transportation Systems*, 24(9):9462–9475.
- Zaidi, K., Milojevic, M. B., Rakocevic, V., Nallanathan, A., and Rajarajan, M. (2016). Host-based intrusion detection for vanets: A statistical approach to rogue node detection. *IEEE Transactions on Vehicular Technology*, 65(8):6703–6714.