Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 08 Issue 01
2024

Research Article
https://doi.org/10.56979/801/2024

Enhancing Vehicular Network Security: An In-Depth analysis of Machine Learning Approaches

Ezzah Fatima[1*], Irshad Ahmed Sumra[2], and Syed Aleem Muzaffar[2]

[1]Department of Information Technology, Lahore Garrison University, Lahore, 54000, Pakistan.
[2]Department of Computer Science, Lahore Garrison University, Lahore, 54000, Pakistan.
Corresponding Author: Ezzah Fatima. Email: ezzahfatima186@gmail.com

Received: October 12, 2024 **Accepted:** December 01, 2024

Abstract: Modern transportation systems heavily rely on vehicular networks, facilitating crucial applications such as autonomous driving, in-car infotainment, traffic management, speed restriction, and road safety. These networks primarily utilize the Vehicular Ad Hoc Network (VANET) architecture, which connects vehicles via roadside units (RSUs) to the edge network and ultimately to a backbone network through wired or wireless connections. However, the open and dynamic nature of VANETs introduces various security challenges that can compromise vehicular communications, potentially jeopardizing the safety and efficiency of intelligent transportation systems. This study examines the current state of security services, common attacks, and application scenarios specific to VANETs, with a focus on machine learning techniques to strengthen these networks. It evaluates advancements, identifies gaps, and suggests future research directions to enhance the robustness and resilience of VANETs in an increasingly connected and automated transportation environment. This study aims to support ongoing efforts to address security issues in VANETs and enable the full potential of vehicular networks in future transportation systems.

Keywords: Vehicular Ad Hoc Network (VANET); Roadside Units (RSUs); Edge Network; Machine Learning Techniques; Vehicular Communications.

1. Introduction

Modern intelligent transportation systems, which use ad hoc networking to improve many aspects of transportation, depend heavily on vehicular networks [1]. Applications including autonomous driving, in-car infotainment, traffic management, speed restriction, and road safety are supported by these networks [1]. More advanced and networked vehicle communications are the result of recent developments, such as the Internet of Things (IoT) and fifth-generation (5G) cellular technology [7]. One of these networks' primary architectures is the Vehicular Ad Hoc Network (VANET). Direct communication between vehicles and roadside infrastructure (V2I) as well as between vehicles themselves is made possible by VANETs [3]. By facilitating the exchange of vital information in real-time, such as traffic conditions and potential hazards, this system helps to improve traffic flow, road safety, and driving efficiency in general. The functionality of VANETs relies on their ability to connect vehicles via roadside units (RSUs) to the edge network, which interfaces with a backbone network through wired or wireless connections. This infrastructure supports critical applications like cooperative collision warning systems and traffic flow optimization [10].

However, the open and dynamic nature of VANETs introduces various security challenges that can compromise vehicular communications [15]. Ensuring the security of VANETs is essential to maintaining the safety and efficiency of intelligent transportation systems. Addressing these security issues is crucial for realizing the full potential of VANETs [16].

This study examines the current state of security services, common attacks, and application scenarios specific to VANETs, focusing on the use of machine learning techniques to strengthen these networks [2]. By evaluating advancements, identifying gaps, and suggesting future research directions, this study aims to support ongoing efforts to enhance the robustness and resilience of VANETs in an increasingly connected and automated transportation environment [12].

2. Architecture of Vehicular Networks

The conventional architecture of vehicular networks is designed to support both autonomous and non-autonomous vehicles, consisting of several key components: On-Board Units (OBUs), Roadside Units (RSUs), cellular base stations (BS), the backbone network, and a Trusted Authority (TA) [13].

2.1. On-Board Units (OBUs)

Installed on vehicles, OBUs are equipped with essential components for sensing various vehicular parameters such as speed, velocity, location coordinates, and proximity to other objects or vehicles [9].

2.2. Roadside Units (RSUs) and Cellular Base Stations (BS)

RSUs and cellular BS serve as interfaces between vehicles and the backbone network. They facilitate the transfer of data from vehicles to the network. Various wireless protocols can be used for this data transfer, with the most common being DSRC (Dedicated Short-Range Communications) and WAVE (Wireless Access in Vehicular Environments) IEEE 802.11p for short-range communication [13]. When a vehicle is out of the range of DSRC/WAVE, a cellular BS is utilized [14].

2.3. Backbone Network

The RSUs connect with different components of the backbone network using wired or wireless connectivity, enabling various types of communications within the network [15].

2.4. Trusted Authority (TA)

The TA is crucial for using basic authorization techniques for vehicles that wish to register within the network, ensuring secure communication and operation [9].

This architecture supports a seamless communication framework within vehicular networks, enhancing the overall efficiency and safety of transportation systems.

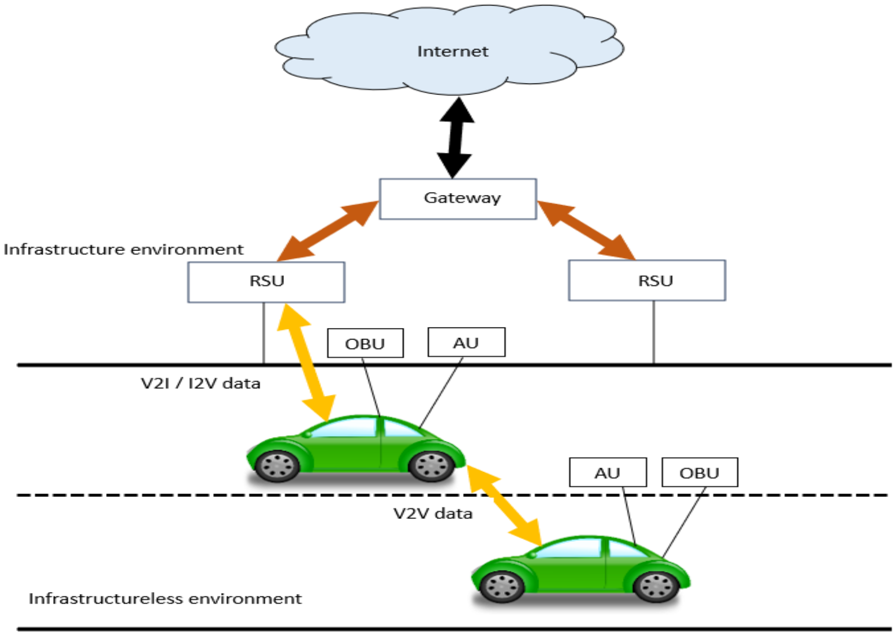

Figure 1. Basic architecture of VANET [14]

3. Vehicular Networks: Evolution and Diversity

With technological advances such as the Internet of Things (IoT) and fifth-generation (5G) cellular technology, vehicular networks have evolved significantly, transitioning from Vehicular Ad Hoc Networks (VANETs) to the Internet of Vehicles (IoV) [7]. IoV enhances network intelligence by integrating an understanding of the environment, including human actions and activities, leading to a new level of communication known as Vehicle-to-Person (V2P) [14].

The emergence of modern technologies such as 5G, Software-Defined Networking (SDN), edge computing, and cloud computing has enabled diverse applications, resulting in the creation of new variants of vehicular networks:

3.1. 5G Vehicular Networks (5GVN):

The integration of 5G technology provides high speed and low latency, essential for applications like fully autonomous vehicles that require data transmission with delays of less than 1ms to make accurate driving decisions [17]. 5G also addresses issues with the traditional VANET wireless protocol IEEE 802.11p, such as intermittent connectivity and insufficient capacity for autonomous vehicles. 5GVNs offer efficient solutions for congestion control, resource sharing, reliability, high throughput, and diverse safety applications.

3.2. Software-Defined Vehicular Networks (SDVN):

SDN simplifies hardware-software management through a flexible networking architecture that effectively handles the dynamic nature of vehicles [12]. SDVN enhances Quality of Service (QoS), routing reliability, and security services [13]. Edge-enabled vehicular network (EEVN) solutions, a variant of SDVN, are ideal for low-delay applications but may face performance degradation due to frequent handovers when many vehicles are connected to a single RSU. Research continues to focus on centralized controller placement to improve scalability and traffic management [15].

3.3. Vehicular Cloud Computing (VCC):

Cloud computing minimizes onboard storage and computation by providing flexible access to virtual services for road users [7]. VCC enables new vehicular network architectures that address issues such as security and privacy through cloud-based solutions. The first cloud-based VANET architecture explored these issues and laid the groundwork for ongoing research in VCC [14].

These advanced variants can be used independently or in combination to enhance scalability, power efficiency, spectrum efficiency, and flexibility in vehicular networks. While these technologies bring significant benefits, they also pose challenges, particularly in ensuring security and privacy. This survey primarily focuses on machine learning-based security solutions for vehicular networks, addressing threats across different levels of communication within these evolving architectures.

4. Security Attacks in Vehicular Networks

4.1. Hardware/Software-based Attacks

Several attacks can take place over the hardware components and software systems of the VANET network to compromise different security requirements. Some of the common attacks are listed below.

Bogus Information: In this type of attack, the attacker sends a piece of bogus or false information to misguide the functioning of hardware or software systems in the vehicles [9].

Timing Attack: This is a side-channel attack. It tries to compromise the cryptographic algorithm of a system by analyzing its timing information which is required to execute the attack [14].

Message Forgery: This type of attack is launched to deceive the recipient about the real sender.

Replay Attack: Also known as a playback attack, where transmission of valid or true data is repeated or delayed maliciously to misguide the functioning of the system [14].

Masquerading Attack: The attacker uses fake identification to gain unauthorized access to the vehicle system [15].

Node Impersonation: The attacker steals the identity of an authorized user to gain access to the system.

GPS Spoofing: This attack tries to fool a global positioning system (GPS) by generating fake signals around the vehicle, misleading the GPS sensor to record fake coordinates and misguide the functioning of the vehicular networks [16].

Tampering Hardware: Deceiving the hardware of a vehicle by providing fake information or creating a fake environment around the vehicle [16].

Routing Attack: Improper functioning of the routing process, which is further classified into different attacks corresponding to malfunctioning at different levels of a network.

4.2. Infrastructure-based Attacks

These attacks infect the system at the infrastructure level:

Repudiation Attack: Takes place at the application layer where a system fails to control the log of actions and tracking of nodes due to malicious manipulations, also known as the act of refusing actions in a system [15].

Platoon Attack: Any action by an attacker to destabilize the functioning of a platoon, where platoon is a concept of grouping vehicles that travel in the same lane with close proximity and similar speed regulations [14].

Session Hijacking: An attacker tries to hijack and get access to the session of data transfer established between the vehicle and destination node [16].

Key-Certificate Replication: The attacker uses duplicate keys and certificates of legitimate users to fool TA and gain access to the network [15].

Unauthorized Access: An attack over the authentication systems of the vehicular networks, compromising authentication parameters such as decrypting login identification (ID) and password for a system account [16].

4.3. Sensor-based Attacks

Sensor-based attacks concern about compromising the authenticity and availability of sensors over the vehicles:

Illusion Attack: An adversary vehicle deceives its own sensors to produce wrong readings and transmits to the network, creating a fake illusion of the scenario over the road to misguide other vehicles and generate false warning messages.

Jamming Attack: The aim of the jammer is to block or interfere with transmission of data from a sensor by sending false alerts or creating a spoofed environment around the sensor.

4.4. Wireless Communication-based Attacks

Attacks over the wireless communication channel:

Brute-Force Attack: An attack over the authentication system of a wireless protocol used for transferring data in V2X communication, where the adversary runs a brute-force algorithm to break access into the medium [15].

Spoofing Attack: The attacker pretends as a legitimate user of the network to gain access over personal information, including spoofing of identification, location, domain-name-server (DNS) information, and internet protocol (IP) address [16].

ID Fingerprinting: Aims to obtain the driver profile and uses it to launch an attack over the system, ensuring correct identification and true profiling of the driver to prevent hacking and theft of the car.

Location Trailing: Illegally getting access to the channel transmitting the vehicle's personal information, allowing an adversary to track the complete path of the target vehicle and follow its location.

Sybil Attack: Creating virtual nodes to launch an attack in a vehicular network, making detection of such virtual nodes difficult.

Denial of Service (DoS) Attack: Launching a bulk of spoofed requests over the server or any other node to make it fully occupied with unnecessary requests and block access to legitimate users.

Spamming Attack: Sending bulk spam messages to consume network bandwidth and increase delay for the transmission of data [15].

Grey-Hole and Black-Hole Attack: Types of wireless routing attacks where a node stops onward forwarding of messages/packets, causing either a complete blackout (black-hole) or partial drop/altered packets (grey-hole).

Man-in-the-Middle Attack: An adversary intercepts communication between two nodes, pretending to be one of them to reply with incorrect information.

Eavesdropping: A sniffing attack where the attacker listens to transmitted information between entities to gain access to personal information without altering or replying to any of the entities.

These categories and attacks provide a comprehensive overview of the security threats faced by vehicular networks, encompassing various aspects from hardware/software vulnerabilities to infrastructure, sensor, and wireless communication-based attacks [16].

5. Security Requirements to Mitigate Attacks

To effectively counter the various security threats targeting vehicular networks, it is crucial to address several key security requirements, including the principles of Confidentiality, Integrity, and Availability (CIA), along with privacy, authentication, and trust.

Confidentiality ensures that sensitive information is accessible only to authorized parties. This can be achieved through robust encryption methods to secure data transmission and storage, and the use of secure communication protocols to prevent unauthorized access and eavesdropping [12].

Integrity guarantees that data remains accurate and unaltered during transmission and storage. This can be accomplished by using cryptographic hash functions and digital signatures to verify the authenticity and integrity of messages, as well as employing error detection and correction mechanisms to prevent and address data tampering [14].

Availability ensures that network services and resources are accessible to authorized users when needed. Implementing robust network architectures and redundancy mechanisms can help resist Denial of Service (DoS) attacks; while monitoring and response systems can quickly detect and mitigate disruptions [16].

Privacy focuses on protecting users' personal information from unauthorized access and misuse. This can be achieved through privacy-preserving techniques such as anonymization and pseudonymization, along with compliance with data protection regulations and standards [15].

Authentication involves verifying the identities of users and devices to prevent unauthorized access. Strengthening authentication processes through multi-factor authentication (MFA) and secure key management practices can protect authentication credentials.

Trust is essential for establishing and maintaining trust relationships between network entities. This can be facilitated by implementing reputation-based systems and trust management frameworks to assess and ensure the reliability of network participants. Continuous monitoring and evaluation of entities' behavior are necessary to detect and address malicious activities [14].

Addressing these security requirements is vital for building a resilient and secure vehicular network capable of withstanding various types of attacks, ensuring safe and reliable operation.

6. Leveraging Machine Learning for Vehicular Security

The proliferation of connected vehicles and smart transportation systems has introduced new security challenges that traditional methods struggle to address. To enhance vehicular network security, researchers have increasingly turned to machine learning (ML) techniques. ML, a branch of Artificial Intelligence (AI) first proposed in 1959 as a self-learning technique for the game of checkers, is now widely explored in almost all areas of networking. It is a computing-based strategy that determines the hidden insights of a dataset without being explicitly programmed, improving performance from its learning experience.

A typical model for traditional ML consists of three phases:

Training Phase: This phase takes raw data and pre-processes it to extract features. The features are input into the ML model to learn patterns and classes of the data [5].

Testing Phase: Here, a new set of data is tested by the ML model for classification based on its learning experience from the training phase [5].

Prediction Phase: Also known as the evaluation phase, this is where the working efficiency of an ML model is evaluated based on quality metrics such as accuracy, false positives, and false negatives. In the case of lower efficiency, the training phase updates its data and/or features to achieve better results [5].

ML techniques are classified into three broad categories: supervised learning, unsupervised learning, and reinforcement learning. Advances in these categories have led to the development of several other learning types, such as deep learning (DL), transfer learning (TL), and federated learning (FL), which work in parallel with the main classes and have received significant attention due to their intelligence in performing various tasks.

By examining the capabilities of ML algorithms, we can explore their effectiveness in identifying and mitigating risks such as unauthorized access, data breaches, and system tampering in vehicular networks. This section highlights key machine learning approaches and their applications in intrusion detection systems, anomaly detection, and secure communication protocols within vehicular environments. Through a comprehensive analysis, the study aims to demonstrate the potential of machine learning in creating robust, adaptive, and intelligent security solutions for the evolving landscape of vehicular networks.

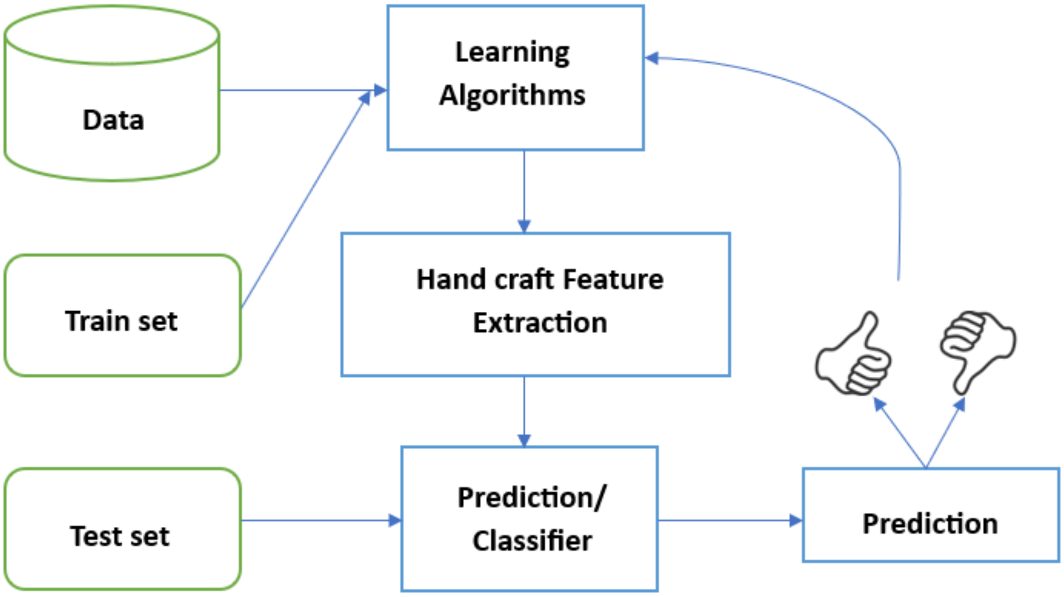

Figure 2. Basic ML Process Flow [15]

6.1. Supervised Learning

Supervised learning is one of the most commonly used ML techniques, where the model is trained on a labeled dataset. This means that each training example is paired with an output label. In the context of vehicular network security, supervised learning algorithms can be used to detect known attack patterns and classify network traffic into normal or malicious categories. Algorithms such as decision trees, support vector machines, and neural networks are particularly effective in these scenarios. By learning from historical data, supervised learning models can make accurate predictions and improve the detection of unauthorized access and other security threats. [2]

6.2. Unsupervised Learning

Unsupervised learning, unlike supervised learning, deals with unlabeled data. The goal is to identify hidden patterns or intrinsic structures in the input data. In vehicular network security, unsupervised learning techniques like clustering and anomaly detection are invaluable for uncovering new, previously unknown threats. Algorithms such as k-means clustering, principal component analysis (PCA), and autoencoders can identify deviations from normal behavior, which may indicate potential security breaches or emerging attack vectors [4]. This approach helps in continuously adapting to new security challenges without the need for labeled datasets.

6.3. Reinforcement Learning

Reinforcement learning (RL) involves training an agent to make a sequence of decisions by rewarding desired behaviors and punishing undesired ones. In vehicular network security, RL can be applied to develop adaptive security protocols that evolve based on real-time feedback from the network environment [5]. For instance, RL can optimize intrusion detection systems by dynamically adjusting detection parameters to minimize false positives and negatives. This approach allows for the creation of more resilient and proactive security mechanisms that can respond effectively to a wide range of cyber threats.

6.4. Federated Learning

Federated learning is a distributed ML approach where models are trained across multiple decentralized devices or servers holding local data samples, without exchanging them. In the context of vehicular network security, federated learning enables vehicles to collaboratively learn a shared prediction model while keeping their data localized. This is particularly useful for preserving privacy and enhancing data security, as sensitive information does not need to be transferred to a central server. Federated learning can improve the overall security of the network by leveraging the collective learning of multiple vehicles, leading to more robust and generalized models [7].

6.5. Transfer Learning

Transfer learning leverages knowledge gained from one task to improve learning in a related, but different, task. This approach is beneficial in vehicular network security, where a model trained on data from one type of vehicle or network configuration can be adapted to another with minimal retraining. Transfer learning reduces the need for large amounts of labeled data and computational resources, making it a cost-effective solution for deploying security models across diverse vehicular environments. By applying pre-trained models to new scenarios, transfer learning facilitates rapid adaptation to emerging threats and evolving network conditions [10].

6.6. Deep Learning

Deep learning, a subset of ML, involves neural networks with many layers (deep neural networks) that can model complex patterns and representations. In vehicular network security, deep learning techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are employed for tasks like intrusion detection, anomaly detection, and behavior analysis. These models can process vast amounts of data and capture intricate relationships, leading to high accuracy in identifying and mitigating security threats. Deep learning's ability to automatically extract features from raw data makes it a powerful tool for enhancing the security of modern vehicular networks [11].

7. Literature Review

The field of vehicular network security has garnered significant attention in recent years, driven by the rapid advancement of connected and autonomous vehicle technologies. This literature review explores various approaches and methodologies developed to safeguard vehicular networks against an array of cyber threats. It delves into the evolution of traditional security mechanisms and examines the burgeoning role of machine learning (ML) in enhancing these systems. By reviewing the latest research, this section aims to provide a comprehensive understanding of the strengths and limitations of different security strategies, particularly focusing on how ML techniques are being employed to detect, prevent, and mitigate cyber-attacks. This review also highlights emerging trends and identifies gaps in the current literature, setting the stage for future research directions in vehicular network security.

Ayoub et al. [1] present an innovative scheme to minimize the invalidity ratio of Vehicular Ad Hoc Network (VANET) packet transmissions while detecting attacks where nodes modify or drop safety messages. The proposed solution evaluates the trustworthiness of data and nodes by combining current and historical behavior, enhancing the detection of unusual traffic patterns. The core of the proposed solution is a four-phase intrusion detection scheme integrating a rule-based security filter, a Dempster–Shafer adder, a node's history database, and a Bayesian learner. Utilizing a clustering mechanism and a rule-based filtering system, the research addresses the challenge of preventing intrusions by malicious nodes in VANETs. This machine learning-driven optimization significantly enhances the security scheme's performance compared to other methods. Extensive simulations validate the effectiveness of the proposed scheme, showing that the fusion of different evidence sources markedly improves the security system's performance. A major contribution of this work is the application of Dempster-Shafer theory to manage uncertainty in VANETs, providing a robust method for handling ambiguous data. This multifaceted approach, combining clustering, rule-based filtering, event-specific trust evaluation, prior knowledge, and a Bayesian learner, effectively reduces the ratio of modified packets and prevents communication disruptions between nodes in VANETs.

Sara & Tomader [2] addresses the inherent problems and vulnerabilities associated with Vehicular Ad Hoc Networks (VANETs) by employing machine learning techniques to enhance network reliability and detect intrusions. The research begins with an overview of VANET networks, exploring the challenges they face and introducing key machine learning concepts relevant to addressing these issues. VANETs facilitate wireless communication between moving vehicles, with the potential to warn drivers of impending hazards. However, like any network, VANETs are susceptible to reliability issues due to various inherent vulnerabilities that threaten their nodes. To mitigate these security risks, the paper investigates the application of machine learning algorithms, which have been extensively studied for their ability to improve VANET reliability and detect potential threats. The study examines various types of attacks targeting VANET networks, underscoring the critical need for effective intrusion detection and accurate risk prediction. By leveraging machine learning algorithms, the authors demonstrate significant improvements in automatic driver efficiency and the early detection of possible dangers. Additionally, the study provides a comparative analysis of different machine learning algorithms applied in VANET networks. This comparison is intended to help readers understand the strengths and limitations of each algorithm in detecting network risks. The algorithms are categorized into supervised, unsupervised, and reinforcement learning types, and are evaluated based on five criteria: detection of outliers, algorithm speed, memory capacity, scalability, and the management of large datasets. Through this comprehensive evaluation, the study offers valuable insights into the most effective machine learning approaches for ensuring the security and reliability of VANETs, thereby contributing to the advancement of secure vehicular communication systems.

Liang et al. [3] explores how machine learning can address the challenges posed by high-mobility vehicular networks, such as rapidly changing wireless channels and diverse Quality of Service (QoS) requirements. It introduces fundamental concepts of machine learning and its various types, including supervised, unsupervised, deep, and reinforcement learning, and discusses their application in vehicular networks. Machine learning techniques can be employed to estimate wireless channel properties, predict traffic flow, and forecast vehicle trajectories. Additionally, the study examines how machine learning can develop decision-making strategies, such as predicting vehicle locations and optimizing routing and scheduling, to minimize congestion and enhance network performance. The study also summarizes open problems and suggest future research areas, providing a comprehensive foundation for further exploration in vehicular network design and optimization. This overview of challenges and machine learning applications makes the research a valuable resource for researchers aiming to improve the performance and reliability of vehicular networks.

Nivedita & Krovi [4] addresses the challenges in Vehicular Ad-Hoc Networks (VANETs), which are vulnerable to various attacks due to their communication and network topology characteristics. To mitigate these risks, the study proposes a novel Hybrid KSVM scheme, combining K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) algorithms, to detect Distributed Denial of Service (DDoS) attacks effectively. The study includes a comprehensive literature review of intrusion detection systems in VANETs using machine learning approaches, comparing the proposed Hybrid KSVM scheme with existing algorithms. The results show that the Hybrid KSVM outperforms others in terms of accuracy, sensitivity, precision, recall, and error rates. The scheme involves training and validating data with SVM, processing input data through nearest neighbor distance calculations, and applying the SVM algorithm to detect misbehaving nodes. Implemented in Python and compared through simulations with other algorithms like KNN, decision tree, and artificial neural network (ANN), the Hybrid KSVM demonstrates superior performance. This study provides valuable insights for researchers, suggesting further exploration of machine learning-based solutions to enhance VANET security.

Arif et al. [5] provide a comprehensive survey on security attacks in Vehicular Ad Hoc Networks (VANETs) published in Vehicular Communications. The study examines the diverse range of security threats and attack scenarios affecting VANETs' communication and applications. Highlighting the dynamic topology and rapid membership changes inherent in VANETs, the authors categorize attacks into communication-based threats compromising data confidentiality, integrity, and availability, and application-specific vulnerabilities targeting systems like traffic management and emergency services. The survey discusses challenges such as efficient key management, secure routing protocols, and robust authentication mechanisms, while evaluating existing countermeasures' efficacy. It also suggests future research directions incorporating technologies like blockchain and artificial intelligence to enhance VANET security, emphasizing the need for resilient measures to safeguard future smart transportation systems.

Rashid et al. [6] present a novel framework for real-time detection of malicious nodes in Vehicular Ad Hoc Networks (VANETs) using machine learning techniques, published in Sensors. The study addresses the critical challenge of ensuring network security in VANETs, where the dynamic and decentralized nature poses vulnerabilities to various attacks. The framework proposes an adaptive approach leveraging machine learning algorithms to identify malicious nodes that can disrupt communication and compromise data integrity within the network. The study discusses the intricacies of VANETs' operational environment and the specific characteristics of malicious behaviors exhibited by nodes, such as denial-of-service attacks and data falsification. By integrating machine learning models for

anomaly detection and classification.

Veres and Moussa [7] provides a comprehensive overview of the transformative role of deep learning (DL) in intelligent transportation systems (ITS). As urbanization accelerates and transportation demands increase, traditional methods for managing traffic and transportation infrastructure are becoming inadequate. This survey highlights the potential of DL in enhancing the efficiency, safety, and sustainability of transportation systems. The authors discuss various DL applications in ITS, such as traffic flow prediction, where convolutional neural networks (CNNs) and recurrent neural networks (RNNs) analyze historical traffic data, weather conditions, and events to optimize traffic management. Additionally, the paper covers DL's role in vehicle detection and classification, traffic signal control, and autonomous driving, emphasizing the significant improvements DL brings to these areas. Veres and Moussa's survey serves as a critical resource for understanding the emerging trends and future directions of DL in ITS, underscoring its importance in addressing contemporary transportation challenges.

8. Comparative Analysis of Existing Work

The table below provides a comparative analysis highlighting the focus, approach, key techniques, strengths, and limitations of each study within the field of vehicular network security.

9. Machine Learning Based Security Solutions for Vehicle Network

In response to the evolving threats faced by vehicular networks, machine learning (ML) offers innovative solutions to enhance security and reliability. Our proposed ML-based security framework integrates several key approaches to mitigate cyber threats effectively:

9.1. Anomaly Detection Using Ensemble Methods:

Implement ensemble learning techniques such as Random Forests or Gradient Boosting to detect anomalous behavior in vehicle communications. By analyzing multiple data sources including vehicle trajectories, communication patterns, and sensor data, this approach enhances the detection of suspicious activities like message spoofing or tampering.

9.2. Behavioral Profiling and Adaptive Authentication:

Develop machine learning models to profile normal behavior patterns of vehicles and drivers. Utilize supervised learning algorithms to establish baseline behavior profiles and adaptive authentication mechanisms that detect deviations indicative of potential attacks, such as unauthorized access attempts or abnormal driving patterns.

9.3. Real-time Threat Intelligence with Deep Learning:

Employ deep learning architectures like Convolutional Neural Networks (CNNs) or Long Short-Term Memory (LSTM) networks to analyze real-time sensor data and detect complex threats such as GPS spoofing or sensor-based attacks. These models can continuously learn and adapt to new attack vectors, ensuring robust threat intelligence.

9.4. Secure Data Fusion and Trust Management:

Utilize Bayesian networks or Dempster-Shafer theory to integrate data from diverse sources (e.g., vehicle sensors, roadside units) while assessing the trustworthiness of incoming information. This approach enhances data fusion reliability and reduces the impact of falsified data injections.

9.5. Dynamic Risk Assessment and Mitigation:

Implement reinforcement learning algorithms to dynamically assess security risks in vehicular networks. By continuously learning from network behavior and attack patterns, these models can autonomously adjust security measures, prioritize responses to imminent threats, and optimize network resources.

9.6. Blockchain-enabled Security for Data Integrity:

Integrate blockchain technology to establish a decentralized and tamper-proof ledger for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. Blockchain ensures data integrity, enhances transaction transparency, and strengthens authentication processes in distributed vehicular networks.

Table 1. Comparative Analysis

Author	Focus	Approach	Key Techniques	Strengths	Limitations
Ayoub et al. [1]	Minimizing invalidity ratio of VANET packet transmissions and detecting attacks	Combines current and historical behavior of nodes	Four-phase intrusion detection (rule-based filter, Dempster–Shafer adder, history database, Bayesian learner)	Improves detection of unusual traffic patterns, reduces modified packets, robust handling of ambiguous data	Complexity of implementation, reliance on historical data
Sara & Tomader [2]	Enhancing network reliability and intrusion detection in VANETs	Machine learning algorithms	Supervised, unsupervised, and reinforcement learning	Comprehensive algorithm comparison, improves automatic driver efficiency, early danger detection	May require significant computational resources, varying performance across different algorithms
Liang et al. [3]	Addressing high-mobility challenges in vehicular networks	Machine learning for network optimization	Supervised, unsupervised, deep, and reinforcement learning	Predicts traffic flow and vehicle trajectories, enhances routing and scheduling	General overview, lacks specific implementation details
Nivedita & Krovi [4]	Detecting DDoS attacks in VANETs	Hybrid machine learning approach	K-Nearest Neighbor (KNN) and Support Vector Machine (SVM)	High accuracy, sensitivity, precision, and recall, reduced error rates	Potential computational complexity, dependence on training data quality
Arif et al. [5]	Survey of security attacks in VANETs	Review of existing threats and countermeasures	Key management, secure routing, authentication mechanisms	Comprehensive threat categorization, future research directions	Focuses on existing technologies, less emphasis on novel solutions
Rashid et al. [6]	Real-time detection of malicious nodes in VANETs	Adaptive machine learning framework	Anomaly detection and classification models	Real-time identification of threats, adaptive to evolving attacks	Implementation complexity, real-time processing requirements
Veres & Moussa [7]	Application of deep learning in ITS	Survey of deep learning techniques	CNNs, RNNs, and their variants	Effective in large-scale data processing, multiple ITS applications	Scalability and deployment challenges, performance metrics evaluation

9.7. Federated Learning for Privacy-Preserving Solutions:

Employ federated learning techniques to train machine learning models across distributed vehicular devices while preserving data privacy. This approach enables collaborative model training without sharing sensitive data, thereby ensuring compliance with privacy regulations and maintaining confidentiality.

These ML-based security solutions collectively enhance the resilience of vehicular networks against a wide range of cyber threats, supporting the safe and efficient operation of connected and autonomous vehicles in modern transportation ecosystems.

10. Critical Analysis of Machine-Learning Based Security Solutions for Vehicular Networks

The table below breaks down the critical components of machine learning-based security solutions for vehicular networks, their limitations, and challenges.

Table 2. Critical Analysis

Key Aspects	Description	Advantages	Challenges and Limitations	Solutions/ Recommendations
Anomaly Detection Using Ensemble Methods	Uses ensemble learning (Random Forests, Gradient Boosting) to detect anomalous behaviors in vehicle communication (e.g., spoofing, tampering).	High detection accuracy obust to noise and outliers. • Effective for diverse attack scenarios.	Requires large datasets for training. High computational cost for real-time processing.	• Improve data diversity. Develop lightweight models for real-time applications.
Behavioral Profiling and Adaptive Authentication	Develops ML models to establish baseline behavior profiles of vehicles and drivers. Detects deviations indicating potential attacks like unauthorized access or abnormal driving.	Accurate detection of unauthorized activities. • Personalized security response based on behavior.	Data privacy concerns when analyzing personal driving data. Challenge in creating comprehensive and accurate profiles.	Implement privacy-preserving methods. • Use federated learning for distributed data training.
Real-time Threat Intelligence with Deep Learning	Utilizes deep learning (CNNs, LSTMs) to analyze real-time data, detecting complex threats like GPS spoofing and sensor attacks.	Ability to detect sophisticated and emerging threats. • Continuous learning and adaptation.	High computational and resource requirements. • Risk of adversarial attacks manipulating learning processes.	Use lightweight deep learning models. • Develop robust models resistant to adversarial attacks.
Secure Data Fusion and Trust Management	Integrates data from multiple sources (vehicle sensors, roadside units) using Bayesian networks and Dempster-Shafer theory to assess data	• Improves reliability of fused data. Reduces impact of falsified data.	Complex integration of heterogeneous data sources. Difficult to maintain trustworthiness in dynamic environments.	• Enhance model robustness with data validation mechanisms. Develop protocols for cross-platform data fusion.

	trustworthiness.			
Dynamic Risk Assessment and Mitigation	Reinforcement learning models autonomously assess risks and adjust security measures based on changing network conditions and attack patterns.	• Dynamic and adaptive to evolving threats. • Optimizes resource usage and threat mitigation.	› High latency due to continuous learning. Requires continuous data flow and interaction.	• Optimize RL algorithms for faster processing. • Develop energy-efficient models for resource-constrained environments.
Blockchain-enabled Security for Data Integrity	Blockchain provides a decentralized, tamper-proof ledger for vehicle-to-vehicle and vehicle-to-infrastructure communications, ensuring data integrity and transparency.	• Strong data integrity and transparency. • Secure decentralized trust model. • Reduced vulnerability to tampering.	• Blockchain may be resource-heavy for real-time applications. Interoperability issues with existing vehicular networks.	• Use lightweight blockchain protocols. • Focus on interoperability with existing systems in vehicular networks.
Federated Learning for Privacy-Preserving Solutions	Federated learning allows model training across distributed devices without sharing sensitive data, preserving privacy.	Ensures user data privacy. › Complies with privacy regulations like GDPR.	• Requires secure communication protocols. › Risk of data leakage through model updates.	› Implement secure aggregation and privacy-preserving techniques. Enhance federated learning protocols for scalability.
Data Quality and Availability	High-quality, diverse datasets are needed for effective ML training in vehicular networks.	Ensures accurate and robust ML models. • Broad data diversity allows comprehensive threat detection.	Difficulty in obtaining diverse, high-quality data across varied conditions. Risk of bias in models due to limited data.	Develop data-sharing initiatives. Use synthetic data for training when real-world data is sparse.
Adversarial Attacks	Adversarial attacks exploit vulnerabilities in ML models, potentially corrupting or manipulating decisions.	• Detects novel threats by analyzing deviations from expected behavior.	Models can be deceived by carefully crafted adversarial inputs. Difficulty in identifying adversarial attacks in real-time.	Develop adversarial training methods. Design robust models that can resist perturbations.
Scalability and Real-time Processing	Real-time processing is necessary for threat detection in large-scale vehicular networks with	• Enables rapid threat detection and mitigation. Supports large-scale, distributed systems.	Scalability issues when processing large volumes of data. Real-time processing may introduce latency.	Optimize algorithms for lower latency. • Employ edge computing for real-time data processing.

	many vehicles and infrastructure components.			
Interoperability and Compatibility	Challenges in integrating ML-based security solutions across diverse hardware platforms, communication protocols, and software systems	Ensures smooth interaction between different systems. • Facilitates widespread adoption.	Compatibility issues with existing infrastructure. • Need for standardization across different technologies.	• Promote standardization efforts. Develop modular ML solutions that can integrate seamlessly with existing systems.
Privacy and Ethical Considerations	ML models may compromise privacy by analyzing sensitive data, creating ethical concerns.	• Ensures user consent and privacy. • Complies with privacy regulations like GDPR.	• Risk of violating privacy if sensitive data is not properly protected. • Ethical dilemmas in data usage.	Implement privacy-preserving models. • Enforce ethical guidelines and ensure transparency in data usage.
Robustness Against Environmental Variability	Vehicular environments are dynamic, including weather, traffic, and road conditions that affect the performance of ML models.	Ensures consistent and reliable security performance. • Adaptable to changing environments.	Models may perform poorly under unpredictable environmental conditions. • Hard to generalize across diverse conditions.	• Develop adaptive models that can adjust to environmental changes. • Improve data collection for varied environmental factors.
Cost and Resource Constraints	ML-based security solutions for vehicular networks are costly, with significant computational and energy requirements.	Potential for high-performance, high-accuracy security systems.	High deployment and operational costs. Resource limitations on edge devices.	• Focus on cost-effective ML models. Optimize for energy-efficient processing in resource-constrained devices.

11. Limitations and Challenges

While machine learning (ML) presents promising avenues for enhancing security in vehicular networks, several challenges and limitations must be addressed to realize their full potential[19,20]:

Data Quality and Availability: ML models heavily rely on high-quality and diverse datasets for training and validation. In vehicular networks, obtaining representative data across various driving conditions, traffic scenarios, and geographical locations can be challenging. Poor data quality or insufficient data diversity may lead to biased models and reduced detection accuracy.

Adversarial Attacks: Adversaries can exploit vulnerabilities in ML models through adversarial attacks, aiming to manipulate or deceive the model's decision-making process. In vehicular networks, adversaries may launch attacks to evade detection systems or corrupt data sources, compromising the reliability and effectiveness of ML-based security solutions.

Scalability and Real-time Processing: Real-time processing requirements in vehicular networks demand ML algorithms that can operate efficiently within strict latency constraints. Ensuring scalability and responsiveness of ML models to handle large volumes of data from numerous connected vehicles and infrastructure components is crucial for timely threat detection and mitigation.

Interoperability and Compatibility: Vehicular networks encompass diverse hardware platforms, communication protocols, and software architectures. Integrating ML-based security solutions across heterogeneous systems requires addressing interoperability challenges and ensuring compatibility with existing infrastructure and communication standards.

Privacy and Ethical Considerations: ML models may inadvertently compromise user privacy by analyzing sensitive vehicle data or personal information. Designing privacy-preserving ML algorithms, enforcing data anonymization techniques, and adhering to ethical guidelines are essential to mitigate privacy risks and maintain user trust in vehicular network security solutions.

Robustness against Environmental Variability: Vehicular environments are subject to dynamic and unpredictable conditions, including weather changes, road conditions, and traffic congestion. ML models must exhibit robustness against environmental variability to maintain reliable performance in diverse operational scenarios.

Cost and Resource Constraints: Implementing ML-based security solutions in vehicular networks entails significant costs associated with data acquisition, model development, training, and deployment. Moreover, the computational and energy requirements of running complex ML algorithms on resource-constrained vehicles and edge computing devices pose additional challenges.

Addressing these limitations and challenges requires interdisciplinary efforts, including advancements in data collection methodologies, robust algorithmic development, secure model deployment strategies, and regulatory frameworks that govern the ethical use of ML in automotive cybersecurity. By overcoming these hurdles, ML-based security solutions can contribute to fostering safer and more resilient connected transportation ecosystems.

12. Conclusion

Vehicular Ad Hoc Networks (VANETs) play a crucial role in intelligent transportation systems, significantly enhancing road safety, traffic flow, and driving efficiency. However, these networks face various security challenges that must be addressed to fully realize their potential. This study reviewed the current state of security services, common attacks, and application scenarios specific to VANETs, with a particular focus on machine learning techniques to strengthen these networks. The key contributions of this study include a detailed description of the architecture and components of vehicular networks, an analysis of different types of security threats and attacks specific to VANETs, an overview of traditional security mechanisms and their limitations, and an in-depth look at the role of machine learning in enhancing vehicular network security. Additionally, a literature review of recent research in this field highlighted strengths and weaknesses, identified gaps, and suggested directions for future research. By addressing the identified security challenges and leveraging machine learning approaches, the robustness and resilience of VANETs can be significantly improved, paving the way for safer and more efficient intelligent transportation systems. Overall, this paper serves as a comprehensive survey of security challenges and solutions in VANETs, with machine learning techniques emerging as a promising area for future exploration.

References

1. Ajaz, F., Naseem, M., Shabaz, M., & Khan, M. A. (2024). An architectural view of VANETs cloud: Its models, services, applications and challenges. *International Journal of Web and Grid Services*, *20*(3), 292-341.
2. Abdallah, A., Alkaabi, A., Alameri, G., Rafique, S. H., Musa, N. S., & Murugan, T. (2024). Cloud Network Anomaly Detection Using Machine and Deep Learning Techniques-Recent Research Advancements. IEEE Access.
3. Multidimensional Security Provision for Secure Communication in Vehicular Ad Hoc Networks Using Hierarchical Structure and End-to-End Authentication - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Basic-architecture-of-VANETs_fig1_327171283 [accessed 9 Jul, 2024]
4. The upsurge of deep learning for computer vision applications - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Basic-machine-learning-process-flow_fig2_338971676 [accessed 9 Jul, 2024]
5. Alsarhan, A., Al-Ghuwairi, A. R., Almalkawi, I. T., Alauthman, M., & Al-Dubai, A. (2021). Machine learning-driven optimization for intrusion detection in smart vehicular networks. Wireless Personal Communications, 117, 3129-3152.
6. Ftaimi, S., & Mazri, T. (2020, March). A comparative study of Machine learning algorithms for VANET networks. In Proceedings of the 3rd International Conference on Networking, Information Systems & Security (pp. 1-8).
7. Liang, L., Ye, H., & Li, G. Y. (2018). Toward intelligent vehicular networks: A machine learning framework. IEEE Internet of Things Journal, 6(1), 124-135.
8. Kadam, N., & Krovi, R. S. (2021). Machine learning approach of hybrid KSVN algorithm to detect DDoS attack in VANET. International Journal of Advanced Computer Science and Applications, 12(7).
9. Arif, M., Wang, G., Bhuiyan, M. Z. A., Wang, T., & Chen, J. (2019). A survey on security attacks in VANETs: Communication, applications and challenges. Vehicular Communications, 19, 100179.
10. Rashid, K., Saeed, Y., Ali, A., Jamil, F., Alkanhel, R., & Muthanna, A. (2023). An adaptive real-time malicious node detection framework using machine learning in vehicular ad-hoc networks (VANETs). Sensors, 23(5), 2594.
11. Veres, M., & Moussa, M. (2019). Deep learning for intelligent transportation systems: A survey of emerging trends. IEEE Transactions on Intelligent transportation systems, 21(8), 3152-3168.
12. Talpur, A., & Gurusamy, M. (2021). Machine learning for security in vehicular networks: A comprehensive survey. IEEE Communications Surveys & Tutorials, 24(1), 346-379.
13. P. Papadimitratos, A. D. La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, "Vehicular communication systems: Enabling technologies,napplications, and future outlook on intelligent transportation," IEEE Communications Magazine, vol. 47, no. 11, pp. 84–95, 2009
14. Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 2, pp. 760–776, Feb 2019
15. M. S. Al-kahtani, "Survey on security attacks in vehicular ad hoc networks (VANETs)," in 2012 6th International Conference on Signal Processing and Communication Systems, 2012, pp. 1–9
16. A. Sumra, I. Ahmad, H. Hasbullah and J. -l. bin Ab Manan, "Classes of attacks in VANET," 2011 Saudi International Electronics, Communications and Photonics Conference (SIECPC), Riyadh, Saudi Arabia, 2011, pp. 1-5, doi: 10.1109/SIECPC.2011.5876939.
17. Shafi, S., Ramzan, S., Sattar, H., Khalid, S., & Hassan, A. (2024). Deep Learning based Smart Healthcare Monitoring System using Sensory Network. Journal of Computing & Biomedical Informatics.
18. Sheikh, Liang, and Wang, "A survey of security services, attacks, and applications for vehicular ad hoc networks (VANETs)," Sensors, vol. 19, no. 16, p. 3589, Aug 2019
19. Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," IEEE Communications Surveys Tutorials, vol. 21, no. 4, pp. 3072–3108, 2019
20. A. Sumra, I. Ahmad, H. Hasbullah and J. -l. bin Ab Manan, "Behavior of attacker and some new possible attacks in Vehicular Ad hoc Network (VANET)," 2011 3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), Budapest, Hungary, 2011, pp. 1-8.
21. Sajjad, R., Khan, M. F., Nawaz, A., Ali, M. T., & Adil, M. (2022). Systematic analysis of ovarian cancer empowered with machine and deep learning: a taxonomy and future challenges. Journal of Computing & Biomedical Informatics, 3(02), 64-87.
22. Siraj, M. A., Rehman, A., Aziz, O., & Khan, M. F. (2021). Systematic Literature Review: Smart Drone for Early Smoke Detection in Forest Using IOT. Journal of Computing & Biomedical Informatics, 2(01), 80-88.
23. I.A. Sumra, Hasbullah, H.B., AbManan, Jl.B. (2015). Attacks on Security Goals (Confidentiality, Integrity, Availability) in VANET: A Survey. In: Laouiti, A., Qayyum, A., Mohamad Saad, M. (eds) Vehicular Ad-hoc

Networks for Smart Cities. Advances in Intelligent Systems and Computing, vol 306. Springer, Singapore. https://doi.org/10.1007/978-981-287-158-9_5.

24. Shah, A. M., Aljubayri, M., Khan, M. F., Alqahtani, J., Sulaiman, A., & Shaikh, A. (2023). ILSM: Incorporated Lightweight Security Model for Improving QOS in WSN. Computer Systems Science & Engineering, 46(2).
25. Abbas, F., Iftikhar, A., Riaz, A., Humayon, M., & Khan, M. F. (2024). Use of Big Data in IoT-Enabled Robotics Manufacturing for Process Optimization. Journal of Computing & Biomedical Informatics, 7(01), 239-248.