This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

Beyond the Attack: Strategies for Attack Endgame Mitigation in VANETs

Mohammed A. Abdelmaguid*, Hossam S. Hassanein*, Mohammad Zulkernine*
* School of Computing, Queen's University, Kingston, Ontario, Canada
Emails:18ma5@queensu.ca, {hossam, mzulker}@cs.queensu.ca

***Abstract*—In Vehicular Ad Hoc Networks (VANETs), neutralizing attack effects is essential for safeguarding human safety and maintaining network efficacy. This paper proposes new strategies to preempt and neutralize the attack endgame. Our strategies specifically address Fake Reporting Attacks, notorious for creating deceptive messages that lead to hazardous and life-threatening situations. The significant frequency and severity of these attacks, culminating in road hazards, underscore the urgency for a concentrated research effort in this field. Our research adopts a dual-faceted approach, comprising both preventative measures for attack endgames and subsequent counteractive strategies. Utilizing the VeReMi for Attack Prediction (VeReMiAP) dataset and leveraging advancements in attack effect prediction, our research introduces a dual-phase approach to mitigate attack endgames. The first phase, endgame prevention, is achieved by deploying advanced Recurrent Neural Network (RNN) models that forecast the potential impact of harmful messages, thereby enabling early intervention. The second phase, endgame countermeasure, employs Geofencing, which uses the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm for strategic message control and Dynamic Blacklist Management (DBM) based on Reputation System-Based Lightweight Message Authentication (RSMA) to exclude malign vehicles. Our findings indicate a reduction in hazardous events by over 40% and a threefold increase in vehicular speed recovery within the network, substantiating the efficacy of our proposed solutions in bolstering VANET security.**

***Index Terms*—Vehicular Ad Hoc Networks (VANETs), Attack Endgame, Mitigation Strategies, Geofencing, Blacklisting, Recurrent Neural Networks (RNNs), Machine Learning (ML), Cooperative Awareness Messages (CAMs), Message Authentication, Road Safety**

I. Introduction

Vehicular Ad Hoc Networks (VANETs) [1] constitute a complex network of mobile vehicles, each outfitted with technology facilitating communication among themselves and roadside units (RSUs). These networks are pivotal in Intelligent Transportation Systems (ITS) [2], offering many advantages such as real-time data exchange, enhanced mobility, scalability, interoperability, and cost-efficiency.

VANETs facilitate robust communication between vehicles. However, they are not immune to security vulnerabilities, including malicious cyber-attacks, which can compromise their reliability and integrity [3]. This research explores a significant threat known as the Fake Reporting Attack [4]. This attack is identified by the fraudulent road hazard alerts, which lead to network disruptions, termed as the Attack Endgame [5].

The security landscape for VANETs can be categorized into two distinct realms: infrastructure-centric and ad hoc-centric, as delineated in Figure 1. For the infrastructure segment, Public Key Infrastructure (PKI) [6] is responsible for authenticating, issuing, and revoking certificates for vehicles seeking to integrate into the network [7]. The ad hoc segment utilizes trust management systems [8] to validate the integrity of information exchanged between authenticated vehicles. Using a dual-structured approach poses significant challenges, particularly when it comes to insider attackers [9], [10] with malicious intent. Such attackers can strike at any point in the system [11]. The proposed trust management system operates subsequent to the PKI-based authentication phase; that is, only messages from vehicles that have already presented valid certificates are evaluated for trustworthiness and potential insider threats. Previous works have highlighted that cryptographic protections alone cannot prevent such misbehavior, since they only validate message origin and integrity but not semantic correctness [12], [13]. Our approach specifically targets this gap by focusing on predicting and mitigating the harmful consequences of fake reporting attacks, thereby complementing the protection already offered by cryptographic standards.

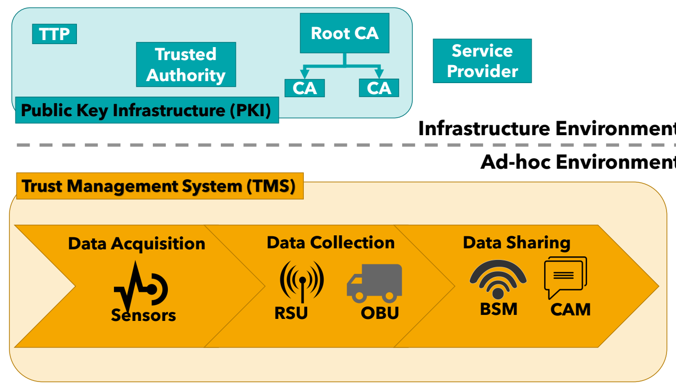

Figure 1: Comparative Overview of Security in VANET Infrastructure Environment and Ad-hoc Environment

Distinct from other studies focusing solely on either pre- or post-attack scenarios [14], our approach encompasses both, as well as addressing the aftermath of attacks within VANETs. We propose a comprehensive, dual-faceted mitigation strategy amalgamating endgame prevention with proactive attack countermeasures. Our strategy begins with endgame prevention, where the utilization of advanced Machine Learning (ML) techniques, especially Recurrent Neural Networks (RNNs) [15], plays a critical role. RNNs are designed for early

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

detection and interception, guarding against communications that may escalate into network-wide disturbances.

Progressing in our goal to secure VANETs, our approach transitions to the attack countermeasure phase, integrating Geofencing [16] and Dynamic Blacklist Management [17]. Employing the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [18], Geofencing orchestrates message flow through cluster management, strategically mitigating network hazards. Dynamic Blacklist Management, supported by Reputation System-Based Lightweight Message Authentication (RSMA) [19], efficiently blacklists suspect vehicles, bolstering both proactive and reactive defenses against these cyber threats. Our system underwent extensive experimental evaluation, encompassing Baseline Performance Comparison, Combined Technique Efficacy, Sensitivity and Threshold Analysis, Interpretation of Modified Reputation Function in RSMA Blacklisting Technique, and Analyzing Mitigation Efficacy Across Diverse Attack Densities. The results indicate a significant reduction in network hazard events and an improvement in vehicular speed recovery.

The existing literature has classified various attacks within VANETs [1]. However, there is a significant research gap in understanding and addressing the far-reaching consequences of these attacks, particularly in terms of their impact on network stability and overall safety. This paper aims to bridge this gap and significantly contributes to enhancing the proactive security of VANET by:

- Conducting an extensive analysis that builds upon previous work [5] by incorporating a comprehensive examination of the cascading effects of attack endgames in VANETs and introducing an algorithmic model designed to predict and mitigate their impact.
- Refining the existing attack mitigation process by integrating strategies to prevent the final stages of an attack and counteract its progression.
- Adopting two critical steps from the established mitigation model: **endgame prevention** to foresee and halt potential threats, and **attack countermeasures** that strategically combine geofencing to contain the attack and blacklisting to prevent its spread.
- Implementing a three-layered geofencing method using DBSCAN clustering and the convex hull algorithm to isolate misbehaving messages within the network.
- Introducing a blacklisting technique derived from RSMA for the removal of misbehaving vehicles.

The subsequent sections of the paper are structured as follows: Section II addresses related work on attack mitigation and misbehavior detection in VANETs. Section III introduces the problem, alongside preliminaries for attack endgame mitigation, and reviews related work on attack prediction in VANETs. Section IV elaborates on the proposed endgame mitigation solution, bifurcated into attack endgame prediction and countermeasure components. Section V describes the adopted experimental setup. Section VI presents the simulation results of the proposed model. Finally, Section VII has concluding remarks and discusses future work.

II. Related work

We emphasize three key areas: (1) work that aligns with our main approach, particularly studies on attack prediction and mitigation; (2) work on detecting misbehaving attacks due to the lack of predictive solutions for these attacks; and (3) work on reputation and trust systems, which are crucial to our proposed solution.

Several studies have leveraged the predictive movement characteristics of vehicles within VANETs for various enhancements, including connectivity and security. For instance, predictive methodologies have been applied to foresee node trajectories, thereby improving network connectivity [20] or facilitating secure communications [21]. Dasgupta et al. [21] devised a strategy for the detection of spoofing attacks predicated on prediction. Their approach utilized a Long Short-Term Memory (LSTM) [22] network applied to the comma2k19 driving dataset [23]. They spotted spoofing attacks where GPS signals are faked to show a vehicle off its course. The LSTM model demonstrated high accuracy in detecting such attacks, with a negligible error rate of mere micrometers and a processing time of 5 milliseconds, comfortably within the VANETs latency standard of 100 milliseconds. Abdelmaguid et al. [5] redefine the attack life cycle in VANETs by incorporating the consequences of cyber-attacks into their predictive modeling framework. By harnessing RNN-based models, including LSTM and Gated recurrent units (GRUs) [24], their study not only predicts the attack endgame with over 80% accuracy within a crucial 3.5-minute timeframe but also emphasizes the significance of understanding the after-effects of such hazards, providing an innovative approach to proactive cyber-attack mitigation in VANET environments. Putra and Sulistyo [25] developed a decentralized, trust-based approach to mitigate Sybil attacks, where attackers create many fake identities to undermine the network. Leveraging the dynamic nature and mobility patterns of vehicular networks, this method independently operates without fixed infrastructure, offering flexibility across various environments. Yet, its effectiveness against sophisticated and adaptive attack techniques needs further exploration.

Numerous studies have concentrated on refining misbehavior detection techniques. Tsukada et al. [26] introduce a framework that utilizes collective vehicular data to pinpoint misbehavior, all while maintaining the privacy of individual nodes. The approach is commendable for considering privacy, an aspect often neglected in misbehavior detection systems. However, investigating system scalability and defense against advanced adversaries mimicking benign behaviors remains an area for further research.

Focusing on reputation systems, Cui et al. [19] propose a framework for 5G-enabled vehicular networks. They introduced a novel reputation-based algorithm named RSMA in their framework to validate messages and enhance network security. Employing elliptic curve cryptography and batch verification, the framework efficiently filters out vehicles with insufficient reputation scores, thereby reducing the presence of untrusted messages. Notably, the framework's design omits the involvement of Road Side Units (RSUs) in the authentication process, significantly reducing computational latency and security risks. Despite its effectiveness in countering common security attacks and maintaining network integrity, the framework's robustness against false positives in the reputation system remains an area for future refinement.

Kang et al. [27] tackle enhancing security in the Internet of Vehicles (IoV) by refining the Delegated Proof-of-Stake (DPoS) mechanism with a focus on selecting trustworthy miners through a reputation-based voting system and an additional step of verification by standby miners. This dual approach aims to ensure data integrity and prevent collusion, positing an increase in the reliability of vehicular data sharing. However, the paper would benefit from a closer examination of potential challenges, particularly regarding the system's adaptability to the dynamic IoV environment and the performance of reputation-based systems across different network scenarios. While the proposed method shows promise for improving IoV security, a more thorough critique and empirical testing are essential to validate its effectiveness and feasibility in real-world applications.

Pseudonym-based reputation mechanisms have emerged as an important complement to cryptographic authentication in VANET security. While cryptographic schemes ensure message integrity and origin, they cannot by themselves prevent insider attacks from vehicles holding valid credentials. Reputation systems address this limitation by evaluating the trustworthiness of vehicles over time, enabling context-aware decisions about identity management and communication. Jaimes et al. introduced a pseudonym change strategy that leverages the reputation of neighboring vehicles to balance anonymity and overhead [28]. Their framework integrates reputation scoring into the timing of pseudonym changes, such that vehicles delay or accelerate changes based on the trustworthiness of immediate neighbors. Beacon messages carry reputation information, and pseudonym changes occur only when both time and reputation thresholds are satisfied. This approach demonstrates how reputation systems can enhance privacy and trust without excessive computational or communication cost.

More recently, Yang et al. proposed Secure Reputation-Based Authentication With Malicious Message Detection, which combines reputation management with authentication to reduce the effects of false or duplicate messages in VANETs, while preserving privacy [29]. The design addresses both lightweight cryptographic operations and trust evaluation, showing that integrating reputation with authentication allows detection of misbehaving actors who might otherwise pass through purely certificate-based schemes.

In the context of trust management systems in VANETs, Hussain et al. [30] provide a comprehensive review of trust management strategies. The study analyzed methods from 2014 to 2019. The authors highlighted the need for future research to focus on space-centric trust approaches and foundational trust establishment, particularly with the growth of the Internet of Vehicles and autonomous technologies. They also stressed the importance of high-quality data for real-time vehicular decision-making and the need for broader datasets to validate predictive trust models. Mahmood et al. [31] propose a distributed trust management system for the Internet of Vehicles (IoV) to enhance misbehavior detection, addressing vulnerabilities that traditional cryptographic methods fail to mitigate, particularly insider attacks. The system evaluates trust using reputation segments based on familiarity, similarity, and timeliness and introduces an intelligent trust threshold for accurately identifying and removing misbehaving vehicles. Their results show that this system outperforms existing models.

III. Problem Statement and Preliminaries

This section covers VANETs' vulnerability to Fake Reporting Attacks, their impact on road safety and vehicle functionality, and the attackers' goals, such as causing hazards and takeovers. It highlights the exploit phase's reliance on deceptive messaging and defines "hazard." Lastly, it examines mitigation techniques, focusing on Geofencing and Dynamic Blacklist Management.

A. Proactive Security Challenges in VANETs

The concept of proactive security in VANETs is fraught with complexities [32]. The unique constraints and inherent nature of VANETs, characterized by high mobility, dynamic topology, and stringent latency requirements [33], increase the challenge of ensuring robust security. Unlike traditional networks, VANETs must accommodate rapid changes and unpredictable patterns, making conventional security approaches ineffective.

Recognizing the importance of pre-empting security breaches [34], our previous work initially focused on predicting the impact of attacks (i.e., attack endgame) in VANETs environments, where we successfully developed predictive models that anticipated the attack endgame. However, prediction alone is insufficient. The true utility of these insights lies in their application towards effective attack mitigation. The transition from understanding the potential impact of an attack to actively countering it is key to improving security in VANET.

The implementation of proactive and prediction-based security solutions is hindered by multifaceted challenges. Firstly, the need for real-time data analysis and threat intelligence strains the limited computational resources inherent to vehicular systems. This limitation poses a significant barrier to executing complex predictive algorithms. In addition, a security system must consider the delicate balance between minimizing false positives to avoid desensitization to security alerts and reducing false negatives to ensure no genuine threats go undetected. Moreover, the necessity for real-time response mechanisms, integral to mitigating potential attacks promptly, underscores the importance of speed in such systems. Integrating these advanced security measures with existing protocols without introducing new vulnerabilities remains a concern.

Lastly, the reliance on collaborative security approaches, which necessitates a harmonized balance between efficient inter-vehicle communication and stringent privacy safeguards, adds another layer of complexity to proactive VANET security.

The next subsections discuss various mitigation techniques and analyze their applicability and effectiveness in the VANETs. We discuss how these adapted techniques can be employed to fortify VANETs against the threats to which they are susceptible.

B. Attack Endgame

This paper addresses the Fake Reporting Attack, which is part of a class of integrity threats in which attackers disseminate counterfeit BSMs or CAMs that misrepresent road conditions. Expanding upon our prior study [5], we concentrate on the attack's terminal—or endgame—stage, where the propagated falsehoods most acutely disrupt traffic flow and jeopardise safety. Various attack endgames have been identified in VANET, including data theft and privacy breaches, network disruption [35], Denial of Service (DoS), and resource drainage [36]. A DoS attack overwhelms a network by flooding it with fake messages, preventing the victim node from accessing the network and disrupting normal operations. In a resource drainage attack, a malicious node depletes resources such as power, bandwidth, and energy by manipulating routing paths and causing excessive consumption, ultimately disrupting the network [37].

Despite their serious implications, there remains a significant gap in addressing attacks that affect the physical state of the network, such as the creation of road hazards [38]. Static thresholds, such as simple bound checks, mitigate similar attacks but fail against semantically valid yet misleading content. Our approach addresses this gap, focusing on content-level misbehavior detection. This oversight is particularly concerning given that such attacks, like the fabrication of fake road hazards, could have life-threatening consequences, potentially more severe than other types of attacks [39]. Recognizing the profound impact of road hazards on both safety and traffic efficiency, our research aims to contribute to the mitigation of these destructive outcomes by focusing on strategies that predict and prevent scenarios leading to serious accidents.

In the context of our research within VANETs, a hazard is defined with two specific quantitative parameters: the speed of vehicles and the number of affected vehicles. A situation is classified as a hazard when the number of vehicles impacted surpasses a predetermined limit, concurrently coupled with a reduction in their speed below a defined minimum. This definition objectively recognizes a hazardous condition on the road. The rationale behind this definition lies in the correlation between high vehicle density and reduced speed, which often indicates potential risks such as traffic congestion or the presence of an obstructive event. By quantifying a hazard in terms of vehicle count and speed, we can accurately identify and respond to situations that pose a threat to road safety. This approach allows for a data-driven assessment of traffic conditions, facilitating timely and effective interventions to mitigate potential dangers in vehicular networks.

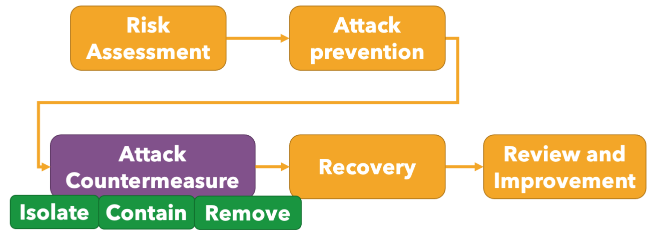

Figure 2: Attack Mitigation Process: A Generic Representation

C. Attack Mitigation

Figure 2 outlines the attack mitigation methodology, including steps to achieve effective outcomes: (1) Risk Assessment, identifying threats and evaluating their impact; (2) Attack Prevention, implementing defensive measures to prevent attacks; (3) Attack Countermeasure, a crucial phase involving activating specific response mechanisms to address and neutralize attacks in progress. This phase is characterized by isolating affected systems, constraining the extent of the attack, and finally removing its causative elements; (4) Recovery, restoring normal operations and repairing damage; and (6) Review and Improvement, conducting post-incident analysis to identify improvements and strengthen the system against future threats.

D. Attack Endgame Mitigation Techniques

In exploring Endgame Mitigation Techniques, our focus on Incident Response Planning (IRP) and Network Segmentation stems from their efficacy in containment and removal within VANETs. IRP, a crucial aspect of our strategy, encompasses steps like monitoring, detecting suspicious activities, and blacklisting, which are effectively implemented through Dynamic Blacklist Management. This approach combines Machine Learning and Rule-based predictions to identify and mitigate threats. Similarly, Network Segmentation, particularly through Geofencing, helps control vehicular communications and mitigate hazard propagation by creating virtual perimeters around high-risk areas. These methodologies not only strengthen our system against malicious threats but also adapt to the dynamic nature of vehicular networks, enhancing overall network security and resilience.

IV. Attack Endgame Mitigation

In this section, we provide an overview of the proposed system, focusing on how prediction and mitigation processes are integrated to effectively reduce the impact of attack endgames. Figure 3 outlines our comprehensive approach to mitigating the attack endgame within vehicular networks. This figure demonstrates the integration of attack endgame prediction with subsequent mitigation techniques, illustrating a robust multifaceted defense mechanism. It highlights the importance of not only attempting to preempt attacks through predictive and preventative measures but also ensuring that, in the event prevention efforts fall short, the system is well-prepared to

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

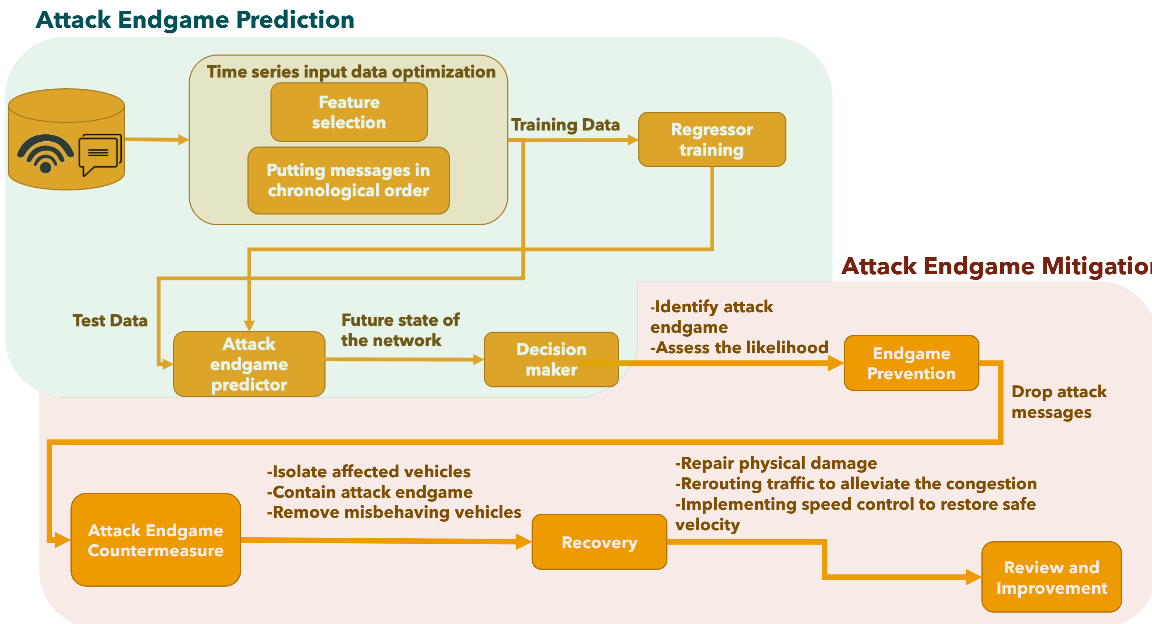

Figure 3: Integration of Attack Endgame Prediction and Mitigation in VANETs

mitigate the repercussions informed by the prediction models. In the subsequent sections, we will explore the prediction and mitigation components in greater detail, elucidating each step illustrated in the figure.

A. Attack Endgame Prediction and Prevention

The top left part of Figure 3 illustrates a workflow designed for predicting the attack endgame in VANETs, starting with the input of exchanged Basic Safety Messages (BSMs) and CAMs. Initially, these messages undergo a time series data preparation process, including feature selection to identify relevant data points and arranging the messages chronologically to preserve the sequence of events. The optimized data is split into two streams: training data for building the predictive model and test data for evaluating the model performance. In the training phase of the predictive model, a regressor is trained on the selected features, fine-tuning the model to predict the future states of the network. For the testing phase, the model, referred to as the attack endgame predictor, takes in new data to predict the endgame of ongoing attacks. Finally, the predictions are fed into a decision-maker component that assesses whether accepting a message will exacerbate the effect of an attack, leading to a decision on how to act on the incoming messages to mitigate potential threats.

1) RNN Architecture: The incorporated RNN architecture is designed with a memory component, enabling the storage of information from prior inputs to influence future inputs and outputs. Given a sequence of inputs $(x_1, x_2, \ldots, x_t)$, the hidden state h_t of the LSTM is computed as:

$$h_t = f_W(h_{t-1}, x_t) \tag{1}$$

where f_W is a function parameterized by weights W, representing the LSTM's cell.

2) Prediction Strategy: We aim to predict the impact of a false traffic alert disseminated by malicious vehicles. The hazard effect, resulting from vehicles acting upon the malicious information, is considered to be the outcome to predict. An LSTM model is used as a baseline to understand the extent of prediction achievable in foreseeing future impacts.

3) Window and Step Sizes: Two significant parameters in our methodology are the window size n and step size r. A window of n messages is considered, and the model predicts the $n+1$ message. The window then shifts by a step size r, and the process repeats. Mathematically, for a given sequence S, the windowing process can be represented as:

$$W_i = (m_i, m_{i+1}, \ldots, m_{i+n-1}) \rightarrow m_{i+n} \tag{2}$$

where W_i is the window of messages, and m represents individual messages in the sequence.

4) Cross-Validation: A cross-validation approach is utilized, incorporating a Train-Validation-Test split technique. Given the temporal nature of the data, temporal component splitting is adopted to maintain the chronological integrity of events. The dataset is divided as 80% for training, 10% for validation, and 10% for testing.

Algorithm 1 RNN Model Architecture for Attack Endgame Prediction

1: Input: Sequence of messages $(m_1, m_2, \ldots, m_t)$
2: LSTM Layer: $h_t = f_W(h_{t-1}, x_t)$
3: Dense Layer: $y = \text{ReLU}(W_h h_t + b)$
4: Output: Predicted impact y

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

5) Model Architecture: Algorithm 1 presents an RNN model architecture that is specifically designed for the prediction of attack endgames. The process initiates with the input of a time-ordered sequence of messages, symbolized as $(m_1, m_2, \ldots, m_t)$. The LSTM layer, a distinctive form of RNN known for its proficiency in managing sequential data, processes each message. It updates its hidden state h_t at every timestep, influenced by the present input x_t and the antecedent hidden state h_{t-1}, following a function f_W parameterized by the set of weights W.

Subsequent to the LSTM layer is a Dense layer, a conventional neural network layer characterized by its fully connected nature, which takes the LSTM's output h_t and applies a Rectified Linear Unit (ReLU) activation. This is mathematically represented as $y = \text{ReLU}(W_h h_t + b)$, where W_h denotes the weights of the Dense layer and b signifies the bias term.

B. Attack Endgame Mitigation

In the bottom right part of Figure 3, we present an extended version of the attack mitigation process depicted in Figure 2, tailored specifically to the context of VANETs and attack endgame. This adaptation aligns with the inherent characteristics and operational dynamics of VANETs. We provide a detailed mapping and explanation of how each step of the attack mitigation process is conceptualized and executed within the VANET environment, focusing on the attack endgame scenario. Attack endgame mitigation in VANETs, primarily constitutes attack endgame prevention and countermeasures. Our approach focuses on specific steps that exhibit potential for targeted mitigation of the attack endgame in VANETs.

1) **Endgame Prevention:** Following the assessment, the next step involves proactive measures to stop the messages causing the attack, effectively preventing the endgame from manifesting.
2) **Attack Endgame Countermeasure:** Should an attack transpire, this phase activates a multi-step countermeasure approach, incorporating substeps:
   - *Isolate:* Segregating affected system components to prevent the proliferation of the attack's impact.
   - *Contain:* Implementing measures to contain the impact of the attack endgame and prevent further compromise.
   - *Remove:* Executing actions to remove misbehaving vehicles from the network to halt the attack.

C. Attack Endgame Countermeasure

In Section III-D, we discussed enhancing the Containment and Removal phases using Geofencing and Dynamic Blacklist Management, respectively. The following section details these methods.

*1) **Geofencing**:* The geofencing mechanism in our system has three key steps:

a. Identify High-Density Regions: Analyze and pinpoint areas with high vehicular concentration prone to hazards.
b. Create Geofences: Set up virtual boundaries around these areas to regulate vehicle communications.
c. Curate Communications: Filter messages from within these zones to prevent hazard spread.

Geofencing is implemented via the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [18] to manage and control the flow of communication within congested areas. By establishing geofenced clusters, the system can isolate and restrict the communication of vehicles within these high-density zones (e.g., clusters). Messages originating from vehicles inside these clusters are selectively dropped, thereby curtailing the spread of harmful information and reducing overall network congestion. Geofenced areas can improve cooperation and information sharing among vehicles inside the cluster while protecting the vehicles outside, enabling a swift collective response to hazards [40].

DBSCAN excels as a clustering algorithm by identifying densely packed data points as clusters, differentiated by areas of lower density, marking a shift from centroid-based clustering methods like K-means to a model capable of uncovering clusters of arbitrary shapes and sizes. This feature is invaluable in real-world scenarios, including vehicular traffic patterns, where data often assume complex, non-convex forms. DBSCAN's renowned proficiency in spatial data analysis aligns well with the dynamic and heterogeneous nature of these patterns, making it especially suitable for applications within such contexts.

The algorithm operates on two primary parameters: *minPts*, the minimum number of points required to form a dense region, which allows the algorithm to be sensitive to the scale of clustering, and ε (epsilon), the maximum distance within which points are considered to be neighbors. This distance parameter ε is pivotal, as it defines the spatial reach of clustering, ensuring that only genuinely adjacent data points are amalgamated into a single cluster.

Figure 4: Generated Traffic Clusters on Road Network Over Time

DBSCAN's resistance to outliers strengthens our geofencing approach by mitigating the impact of anomalous data points, making it ideal for VANET security. Its use of Euclidean distance supports robust geofencing mechanisms, filtering communications from high-risk zones and enhancing network resilience. The Convex Hull algorithm further refines this approach by creating efficient geofences that adapt to real-time changes, ensuring secure management of vehicle clusters.

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

Figure 4 illustrates how clusters form over time in different areas on the map. At some point in time, a single or multiple clusters might form, as seen on the map. The implementation of Geofencing unfolds in four steps:

Algorithm 2 Geofencing Using DBSCAN Clustering

1: **Inputs:**
2: df_Test: DataFrame with 'x_pos', 'y_pos'
3: eps: Max distance for neighborhood consideration
4: $min\_samples$: Min samples for a core point
5: $threshold$: Min vehicles for a high-density cluster
6: $near\_distance$: Max distance to be near high-density
7: **procedure** DBSCAN_GEOFENCING(df_Test, eps, $min\_samples$, $threshold$, $near\_distance$)
8:     Initialize $dbscan$ with eps, $min\_samples$
9:     Assign clusters to df_Test using $dbscan$
10:     Extract high-density clusters
11:     Mark vehicles in/out of high-density clusters
12:     **for all** $vehicles$ not in high-density clusters **do**
13:         **if** $vehicle$ on a road that has high-density clusters **then**
14:             Calculate $distance$ from $vehicle$ to all vehicles in high-density cluster
15:             **if** $distance \leq near\_distance$ **then**
16:                 Mark $vehicle$ as 'Near Hazard'
17:             **end if**
18:         **end if**
19:     **end for**
20:     **for all** $cluster\_id$ in $high\_density\_clusters$ **do**
21:         Calculate Convex Hull for the cluster
22:     **end for**
23: **end procedure**

1) Identify Clusters: Use DBSCAN for finding high-density areas in the BSMs and CAMs dataset as potential clusters.
2) Construct Geofence Boundaries: Apply the Convex Hull algorithm to outline geofence boundaries for each cluster, forming a minimal enclosing polygon.
3) Implementing the Algorithm: Utilize well-established methods (e.g., Graham's scan) for Convex Hull implementation, sorting points by polar angle and constructing the hull.
4) Execute Convex Hull Algorithm: Perform the Convex Hull algorithm on cluster points to create the polygonal geofence.

Algorithm 2 outlines a method to apply geofencing in VANETs using the DBSCAN algorithm. It begins by setting key parameters, including vehicle positions, the maximum neighbor distance (eps), minimum cluster size (min_samples), and criteria for identifying high-density clusters (lines 1 - 6). DBSCAN then groups vehicles based on proximity, marking those in (line 14) or near (line 16) high-density clusters. The algorithm also uses Convex Hulls to outline these clusters, providing visual boundaries (line 21).

*2) **Dynamic Blacklist Management (DBM)**:* DBM is a critical component for implementing the Removal step. Adopting the reputation algorithm, intricately woven into the Reputation System-Based Lightweight Message Authentication (RSMA) algorithm, is a pivotal cornerstone in this implementation.

A confluence of strategic advantages underscores the selection of the RSMA algorithm. Primarily, its intrinsic compatibility with CAMs and the centrality of authority mechanisms bolster its applicability. Furthermore, the algorithm considers the malicious message's consequential effects. Augmented by its advanced developmental maturity and enhanced scholarly recognition, the reputation algorithm of RSMA emerges as an astute choice for Dynamic Blacklist Management.

In our implementation, several key modifications have been integrated into the RSMA algorithm to enhance its adaptability and efficacy within DBM. Notably, we eliminated the constraint of time windows and introduced a novel definition of the "Message Effect". Additionally, feedback participation has been broadened, allowing a more diverse range of vehicular entities to contribute to the reputation assessment.

The modified RSMA algorithm (Algorithm 3) is implemented through a structured process that extends the original RSMA framework. Each procedure contributes to the final calculation of a vehicle's reputation score, which is essential for Dynamic Blacklist Management (DBM). The process begins with the *Effect Degree* step (lines 8–17), where the impact of a vehicle's behavior is quantified by multiplying the number of feedback reports with the appropriate message effect score, determined according to feedback percentage. This ensures that the severity of misbehavior reflects both the occurrence and its spread in the network. The next stage, *Objective Evaluation* (lines 20–23), aggregates reputation scores from neighboring vehicles, incorporating not only direct observations but also second-hand feedback from trusted peers. This broadens the evaluation and improves resilience against insider threats. The *Historical Reputation* procedure (line 26) then incorporates past behavioral data into the assessment, maintaining continuity and preventing abrupt resets of reputation values. Finally, the *ReputationScore* procedure (lines 29–33) combines all components—Effect Degree, Objective Evaluation, and Historical Reputation—using weighted coefficients (α, β, γ) to compute an adaptive reputation score.

The resulting reputation value serves as a dynamic measure of vehicle trustworthiness within the VANET. It integrates first-hand information with second-hand feedback, applying a weighted methodology that balances both perspectives. Unlike the original RSMA, the modified design continuously updates the score in real time, reflecting evolving behavior more accurately and ensuring that malicious activity is detected promptly. These refinements allow the reputation system to better capture current network dynamics, thus enhancing DBM's ability to mitigate misbehavior and safeguard communication reliability.

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

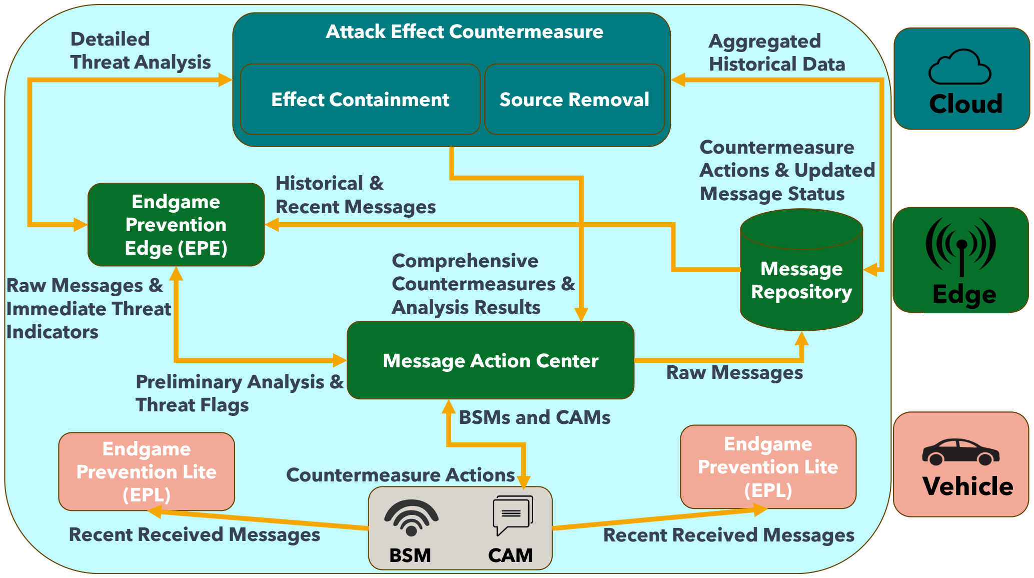

Figure 5: VANET Security Architecture for Attack Endgame Prediction and Mitigation

D. Proactive and Retroactive Mitigation of Attack Endgames

Our system model follows a standard VANET architecture consisting of vehicles with On-Board Units (OBUs), Roadside Units (RSUs), and a trusted authority. Communication is supported by IEEE 802.11p/ITS-G5 for short-range V2V and 3GPP C-V2X (PC5 Mode 4) for broader coverage. Security relies on IEEE 1609.2, which provides authentication and integrity through digital signatures. Within this secured environment, our work complements cryptographic protection by mitigating residual risks of fake reporting attacks using predictive endgame prevention and dynamic blacklist management.

Figure 5 presents a detailed view of the main components of our system, how they are interconnected, and where they are located within the network. Our system includes Endgame Prevention Lite (EPL), a simplified version of the endgame prevention system. EPL is integrated into vehicles to enable early detection and rapid response mechanisms. The BSMs and CAMs processed by the EPL are transmitted to the Message Action Center at the same time. This center subsequently forwards these messages to both Endgame Prevention Edge (EPE) and Message Repository while also orchestrating the dissemination of countermeasure directives back to the vehicles.

The Endgame Prevention model operating at the edge possesses enhanced data access and extended processing time, allowing for a more thorough analysis. The refined data obtained are then relayed to the Attack Countermeasure System, tasked with addressing the anticipated attack endgames. This system also communicates its findings back to the Message Action Center, which plays a pivotal role in determining whether to discard messages identified as part of the attack.

Furthermore, the Message Repository continuously accumulates messages, contributing additional data to the Attack Effect Countermeasure for use in the mitigation process. It also retains updated statistics on message interactions and the specific countermeasures employed for each message, serving as a vital resource for future predictive analysis.

In the final stage of the process, the Attack Effect Countermeasure, upon reviewing the insights gleaned from both the EPE and the Message Repository, executes strategic actions such as geofencing or blacklisting, targeting specific areas and vehicles. These decisions are then conveyed back to the Message Action Center, completing the system's comprehensive approach to threat mitigation.

We use a train-once, infer-many approach. Models are trained offline using diverse traffic and attack scenarios, then shared with vehicles and RSUs for lightweight, real-time inference. This avoids the need for re-training in each road setting. Detection relies on message content and spatiotemporal context, while DBSCAN-based geofencing and dynamic blacklist management handle containment and the exclusion of misbehaving vehicles, including registered insiders. This design follows recent VANET misbehavior-detection work showing that efficient, edge-level inference can adapt to changing traffic conditions and that reputation-based methods strengthen cryptographic protections with behavior-focused trust [41], [42].

The proposed system remains computationally efficient, with inference and clustering well within OBU/RSU capabilities. This is consistent with recent work showing that ML-based misbehavior detection is feasible at the vehicular edge [43]. Additionally, our approach's reliance on lightweight prediction and clustering matches the kinds of trade-offs shown in recent works, which deliver high detection performance without prohibitive resource requirements. Therefore, our design assumes feasibility even under realistic OBU/RSU hardware constraints [44].

V. Experimental Configuration

In this section, the experimental configuration is presented, outlining the dataset, simulation scenarios, and communication

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

Algorithm 3 Dynamic Blacklist Management: Reputation Score Calculation

Inputs:
threshold_feedback (Range: 10 to 90)
effectScore_low (Range: 0 to 1)
effectScore_mid (Range: 0 to 1)
effectScore_high (Range: 0 to 1)
Ensure: effectScore_low + effectScore_mid + effectScore_high = 1
procedure EFFECTDEGREE($targetVehicle$)
  feedbackPct ← feedbackPctDic.get($targetVehicle$, None)
  if feedbackPct ≥ threshold_feedback **then**
    effectScore ← effectScore_high
  else if threshold_feedback/2 ≤ feedbackPct < threshold_feedback **then**
    effectScore ← effectScore_mid
  else
    effectScore ← effectScore_low
  end if
  N ← feedbackVehicleDic.get($targetVehicle, \{\}$) .get($'count', 0$)
  return effectScore $\cdot N$
end procedure
procedure OBJECTIVEEVALUATION($targetVehicle$)
  feedbackVehicles ← feedbackVehicleDic.get($targetVehicle, \{\}$) .get($'vehicles', []$)
  reputationScoreList ← [reputationScoreDic.get($vehicle, 0$) for vehicle in feedbackVehicles]
  Calculate E based on reputationScoreList
  return E
end procedure
procedure HISTORICALREPUTATION($targetVehicle$)
  return reputationScoreDic.get($targetVehicle$)
end procedure
procedure REPUTATIONSCORE($targetVehicle, \alpha, \beta, \gamma$)
  D ← effectDegree($targetVehicle$)
  E ← objectiveEvaluation($targetVehicle$)
  H ← historicalReputation($targetVehicle$)
  $RS\_new \leftarrow \alpha \cdot H + \beta \cdot E + \gamma \cdot D$
  return RS_new
end procedure

technology employed to evaluate the proposed approach.

A. Adopted Dataset

The work presented in this paper utilizes the VeReMiAP dataset [45], a detailed collection of vehicular communication data accessible from Borealis [46], 2 gigabytes in size, and encoded in JSON format. This dataset chronicles the sequence of messages exchanged between vehicles, with BSMs and CAMs logged at one-second intervals. Each entry, representing either a BSM or CAM, includes standard data fields, with two additional fields for our analysis. The Fake Reporting Attack field signifies messages that falsely report incidents, while the attack effect field records the timing of resultant hazard events, independent of the timestamps of the attack messages. Figure 6 shows a screenshot from the OMNET++ simulation environment, illustrating the LuST scenario, which features a map of Luxembourg highlighted in red. The SUMO traffic model is also integrated, depicted by the movement of vehicles on the map, representing real-time traffic dynamics within the simulation.

B. Simulation Scenario

The Luxembourg SUMO Traffic (LuST) Scenario [47] was selected for its high-fidelity representation of a VANET environment, leveraging real-world vehicular and traffic data. This scenario includes a varied mix of urban elements such as intersections, roundabouts, and highways, providing a rich context for VANET research.

Attack Model: For the attack model, we focus on the Fake Reporting Attack. This Attack is particularly dangerous as it directly impacts road safety and can create hazardous traffic conditions. The adversarial model in this study assumes that malicious entities are legitimate VANET participants holding valid cryptographic credentials. This means that although their messages comply with IEEE 1609.2 requirements for authentication and integrity, the reported content may still be false. For example, an attacker can generate fake Cooperative Awareness Messages (CAMs) or Basic Safety Messages (BSMs) that are properly signed but misrepresent their status or road conditions, thereby creating false hazard alerts.

The simulated action in our scenario includes vehicles slowing down, which may create a hazard if the number of affected vehicles exceeds a certain limit. A hazard, in our context, occurs when three conditions are met: (1) The vehicle is on the reported road; (2) The vehicle's speed drops below 20% of the road's maximum speed; (3) The number of affected vehicles exceeds a certain threshold. More detailed information about this attack is presented in our previously published work [5].

C. Communication Technology

The VANET presented in this work is defined according to the F2MD framework built on VEINS (OMNeT++ + SUMO), the same model underpinning the VeReMi dataset. The connection technology we adopt in this study is Dedicated Short-Range Communications (DSRC). F2MD implements both dominant V2X access technologies: ITS-G5 (IEEE 802.11p), which delivers sub-10 ms, one-hop Dedicated Short-Range Communications (DSRC), and C-V2X PC5 Mode 4, which uses LTE/5G sidelink scheduling to achieve comparable (¡20 ms) latency with higher link reliability. Mobility traces are generated over urban-grid topologies identical to those in the original VeReMi scenarios, ensuring realistic vehicle density, message rates, and channel contention. This dual-stack, widely accepted setup provides a well-established benchmark for evaluating security countermeasures in VANETs. Although

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

simulations used DSRC, the proposed approach is not tied to this medium. Since CAMs and BSMs are standardized across ITS-G5 and C-V2X, the proposed solution remains applicable in 5G-based environments. This portability aligns with findings that higher-layer misbehavior detection remains essential regardless of the underlying PHY/MAC [48], [49].

Despite the evolution of cellular technologies that offer extensive coverage and cater to non-emergency functionalities such as entertainment systems, DSRC remains prominent in the VANET security landscape [50]. This is attributed to its steadfast reliability, the ability for direct communication bypassing the need for cellular infrastructure, and the dedicated spectrum allocation crucial for the necessities of real-time, safety-oriented vehicular communication.

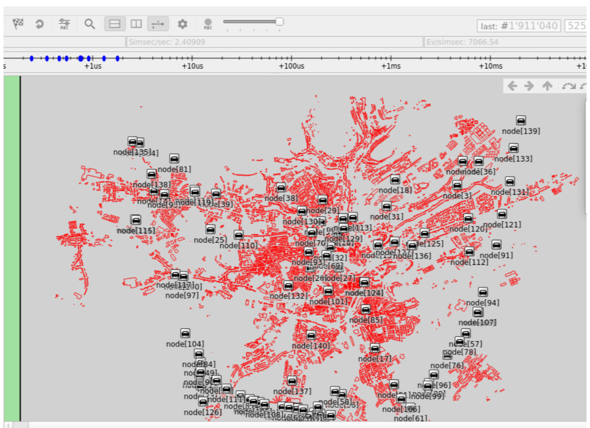

Figure 6: Simulation in OMNET++: LuST Scenario with SUMO Traffic

VI. Experiments and Results

We conducted our simulation using Instant Veins 5.2, a preconfigured Linux instance with Veins [51]. SUMO is used for traffic simulation [52], while OMNeT++ and Veins handle network simulation, providing the necessary infrastructure and tools [51]. The scenario parameters are outlined in Table I. Beaconing rate in Table I is set to 1 Hz, which follows the configuration of the VeReMiAP dataset used in our experiments. While real-world safety-critical applications such as BSMs and CAMs typically operate at 10 Hz, both 1 Hz and 10 Hz have been employed in VANET studies, depending on the objectives and computational constraints of the experiments [53]. The 1 Hz setting is widely adopted in misbehavior detection research. The speed values in the table represent maximum allowed speed limits within the LuST scenario, while the instantaneous vehicle speeds vary dynamically throughout the simulation. These choices maintain consistency with prior VANET simulation frameworks while preserving the realism needed to evaluate attack endgames and mitigation strategies.

We present a series of five experiments, each designed to evaluate and validate different aspects of the proposed mitigation strategies within VANETs. The five experiments are as follows,

1) Baseline comparisons serve as a foundational assessment, establishing a benchmark against which the effectiveness of the proposed techniques can be measured.
2) Combinatorial approaches examine the combined effects of various mitigation methods, exploring how these integrations enhance overall system resilience.
3) Sensitivity and threshold experiments are then conducted to understand the system's robustness under varying operational parameters, offering insights into optimal configurations.
4) Interpretation of the modified reputation function in RSMA blacklisting technique and an in-depth analysis of the tailored approach to Dynamic Blacklist Management.
5) Analyzing mitigation efficacy across diverse attack densities investigates how the system performs under different levels of adversarial pressure, assessing its adaptability and effectiveness in varied attack scenarios.

Parameter	Value
Mobility	SUMO
Mobility scenario	City Streets
Simulation area	2300m x 5400m
Simulation duration	1500s - 4000s
Threat	Fake Reporting Attack
Attack probability	0.05 - 0.5
Number of Vehicles	200 - 850
Exchanged messages	BSMs, CAMs
MAC implementation	802.11p
Beaconing Rate	1 Hz
Speed	Maximum Speed Limit

Table I: Simulation Parameters

A. Baseline Comparisons

We chose our baseline approach due to the notable lack of existing methods that specifically target attack endgame prediction in VANETs. This gap made it essential to establish a baseline for measuring the effectiveness of our attack endgame mitigation, which fundamentally relies on accurate endgame predictions. Additionally, previous research on the dataset we used, as well as related datasets like VeReMi, has primarily focused on the detection of attacks rather than on prediction or the attack effect. Therefore, our baseline fills this research gap and provides a necessary benchmark for evaluating how well our mitigation strategies perform when grounded in predictive modeling.

As depicted in Figures 7 and 8, the Endgame Prevention technique exhibited the most substantial decrease in hazard events, reducing the count to 24,361, a remarkable 37% reduction. This significant decrease highlights the robustness of employing machine learning models for the predictive prevention of hazardous scenarios.

Upon analysis, it was observed that in the State Under Attack scenario with no mitigation techniques applied, 38,660 hazard events were recorded. The implementation of the Blacklisting technique marked a significant decrease in these events to 33,970, indicating a 12% reduction. This finding not only underscores the immediate efficacy of the Blacklisting

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

approach but also highlights its potential in reducing hazard events.

Applying the Geofencing technique further contributed to hazard mitigation, resulting in a decrease to 32,203 hazard events. This represents a 17% reduction compared to the original state, attesting to the effectiveness of this technique, particularly in its capacity to cluster and regulate messages in clustered areas.

The preliminary experiment shows some superiority of prevention against mitigation. This apparent superiority of the preventative approach can be ascribed to the temporal efficiency of machine learning-based methods utilized in Endgame Prevention, as compared to the more time-intensive nature of mitigation techniques. Specifically, the latency inherent in mitigation processes is highlighted by the duration required for the reputation function to accumulate sufficient data for the blacklisting of malicious vehicles and the time needed to establish geofenced areas. Such delays in the mitigation response could extend the overall duration required to neutralize threats, thereby underscoring the relative advantage of the preventative strategy in this experimental framework.

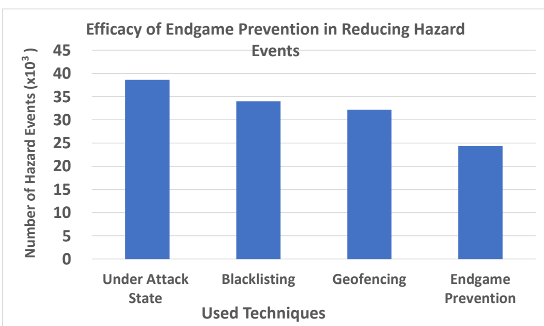

Figure 7: Impact of Endgame Prevention Technique on Hazard Event Reduction

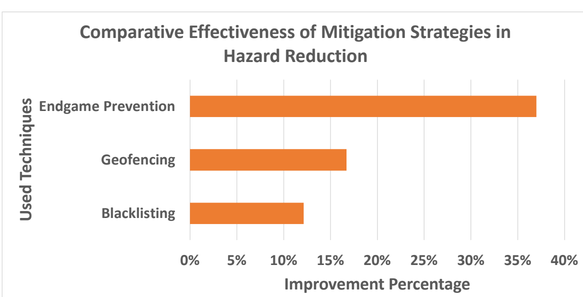

Figure 8: Comparative Analysis of Hazard Event Reduction Percentages by Mitigation Technique

B. Combinatorial Approaches

We present two distinct types of experiments. The first set of experiments investigates the synergistic effects of combining various attack endgame mitigation techniques. The second set focuses on integrating a singular attack endgame mitigation technique with attack prevention measures. Both experimental approaches aim to discern the efficacy of these techniques in isolation and in concert, thereby providing insights into the optimal configurations for VANET security enhancement.

1) Evaluations of Combined Mitigation Techniques: The experiment on Combinatorial Approaches investigated the synergistic impacts arising from the integration of diverse mitigation techniques, shedding light on the enhanced efficacy derived from their combined application in addressing hazard events.

Figure 9 presents a comparison of different methods used in VANETs to reduce hazards. It shows the percentage increase in hazard reduction achieved when various mitigation strategies are used together.

The initial findings indicate that combining Blacklisting (referred to as BL in the figure) with Endgame Prevention culminates in a substantial reduction of hazard events, with a recorded decrease of 49%. Meanwhile, using Geofencing (noted as Geo in the figure) alongside Endgame Prevention achieves a 41% reduction in hazard events. This result highlights the effective interplay between the proactive cessation of potentially hazardous messages via machine learning predictions constituting Endgame Prevention and the active dismissal of communications from confirmed malevolent sources using Blacklisting.

Conversely, when Endgame Prevention is excluded from the strategy and Blacklisting is merged with Geofencing, the outcome manifests as a 24% decrease in hazard events. This is an improvement, though it is less substantial than other combined techniques. This difference highlights that using the Endgame Prevention technique alongside other strategies significantly boosts the effectiveness of the mitigation outcome.

Ultimately, these combinatorial insights highlight the value of deploying a multi-angled approach to mitigation. Individual techniques carry specific merits, yet their collective deployment reveals augmented dimensions of defense against hazards, emphasizing the strategic advantage of a multi-faceted mitigation framework.

2) Augmentation with Endgame Prevention: The findings derived from this experiment underscore the prominence of endgame prevention, prompting a dedicated experimental analysis to assess the augmentative potential of various countermeasure methods when integrated with endgame prevention. This evaluation provides insights into the extent to which the predictive faculties of Endgame Prevention can bolster the efficacy of other mitigation techniques. Such an analysis is not only instrumental in enhancing the overall security framework but also pivotal in understanding the synergistic interactions between different defensive strategies.

The analysis, as illustrated in Figure 10, reveals a consistent pattern: Endgame Prevention invariably boosts the effectiveness of other countermeasures. Notably, its combination with Blacklisting elevates the hazard reduction rate from 17% (with Blacklisting alone) to an impressive 49%. This substantial increase highlights the synergistic relationship between the two methods.

Similarly, combining Geofencing with Endgame Prevention also demonstrates a remarkable improvement. Geofencing

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

alone accounts for a 12% reduction in hazards, but this figure ascends to 41% as indicated in Figure 10, when coupled with Endgame Prevention. The strategic approach of Geofencing, involving the clustering and regulation of messages in designated suspicious zones, is amplified by the insight provided by Endgame Prevention.

When Blacklisting and Geofencing security techniques are used together, they can reduce hazards by 24%. When Endgame Prevention is added to this combination, the hazard reduction increases to 35%. Although this improvement is good, it is not as significant as the improvement seen in other combinations, which suggests that there may be some overlap when all three techniques are used together.

In conclusion, these findings underline the transformative impact of Endgame Prevention. While it is effective on its own, its value lies in enhancing the performance of other countermeasures, strengthening the overall framework for mitigating potential hazards.

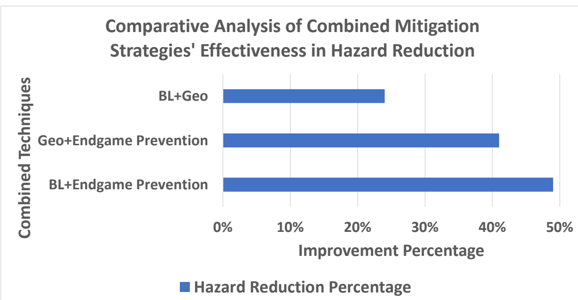

Figure 9: Enhancement Ratios in Hazard Mitigation through Combined Mitigation Strategies

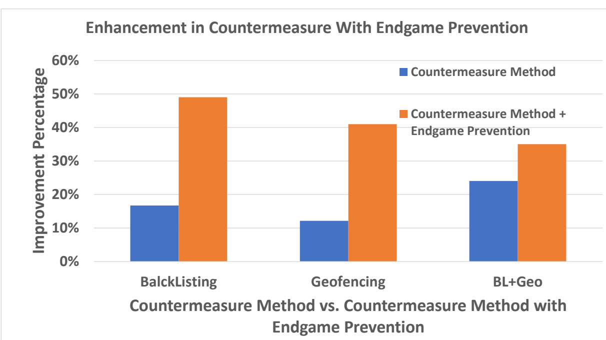

Figure 10: Synergistic Efficacy of Endgame Prevention in Combination with the Different Mitigation Techniques for Hazard Reduction

C. Sensitivity and Threshold Analysis

To ensure that our model is effective and adaptable, it is crucial to evaluate its performance across various settings. This section focuses on examining how well the system adjusts to changes in threshold values, especially in the areas of clustering and blacklisting. By altering these threshold levels, we aim to test the model's stability and determine the best settings for its key parameters. This approach allows us to understand how the model responds under various conditions and guides us in fine-tuning it for optimal performance.

1) Impact of Geofencing Parameters: Min Core and Epsilon Analysis: This experiment aimed to optimize the geofencing technique, which uses the DBSCAN clustering algorithm, by examining its main parameters: Min Core and Epsilon (Eps Value). Min Core is the minimum number of points needed for a dense cluster, while Epsilon is the maximum distance for considering points as neighbors.

Figure 11 shows the inverse correlation between the number of hazardous events and the increment in the Epsilon value, which correlates with an expanded geofenced area. The number of hazard events consistently decreases with these parameter adjustments. Initially, with an Epsilon of 0.2 and a Min Core of 20, the number of hazardous events is 38,157; this decreases to 32,203 as the Epsilon is raised to 0.8 and the Min Core is lowered to 7.

Table II shows an escalation of the Epsilon value from 0.2 to 0.8 results in a proliferation of clusters, signifying a more granular division of the area. Conversely, a reduction in the Min Core Points from 20 to 7 indicates a decreased threshold for cluster formation, albeit over a more extensive area due to the increased Epsilon. The data from the table further reveal that an increase in Epsilon and a decrease in Min Core Points correspond with an augmented impact of geofencing on the vehicular network, evidenced by an increase in affected vehicles from 381 to 743 and a prolongation of geofencing activity from 444 to 752 seconds. This demonstrates that alterations in these parameters not only extend the coverage area but also the duration of geofencing, which might not be optimal under certain scenarios, necessitates careful consideration. Hence, the selection of an optimal combination of Epsilon and Min Core Points is relegated to the discretion of the administrator, who must weigh the acceptability of the increased geofencing duration against the efficacy in hazard reduction.

2) Adaptive Thresholding in Dynamic Blacklist Management: Hazard Event Analysis: In the process of Blacklisting, setting the right threshold is key to determining when a sender should be marked as malicious. Our study explores how different threshold settings impact the system's success in reducing hazard events.

Figure 12 shows that starting with a threshold of 0.2, which is quite forgiving, leads to 38,660 hazard events. This high number indicates that many harmful messages are still getting through because the threshold is not strict enough. However, when we increase the threshold to 0.4, there is a noticeable decrease in hazard events to 33,970, suggesting that a higher threshold is more effective at reducing hazard events.

Raising the threshold to 0.7 and 0.8 does not reduce the number of hazard events; it remains at 33,970. This finding suggests that once the threshold reaches a certain level (here, 0.4), raising the threshold does not increase the number of hazards.

This experiment highlights the importance of choosing the blacklisting threshold carefully. While a threshold that is very low can let too many risks through, setting it extremely high does not necessarily bring extra benefits. The data indicates

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2025.3624720

that a threshold around 0.4 is optimal for effectively minimizing hazard events.

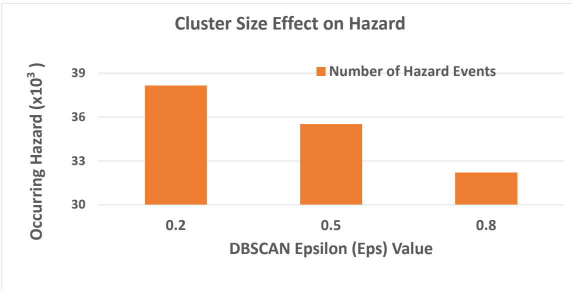

Figure 11: Correlation Between Epsilon and Min Core Values on Cluster Formation and Area Coverage in Geofencing

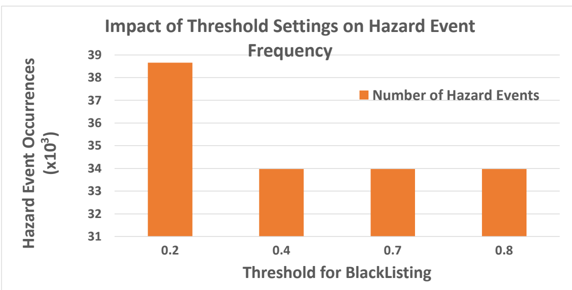

Figure 12: Effect of Threshold Adjustment on Hazard Event Frequency in Blacklisting Strategy

Eps Value	Min Core Points	Number of Geofenced Areas	Number of Affected Vehicles	Total Duration Time for Geofenced Areas	Hazard
0.2	20	9	381	44	38157
0.5	12	21	706	709	35516
0.8	7	24	743	752	32203

Table II: Impact of Parameter Adjustments on Geofencing Efficacy and Duration on Hazard Events

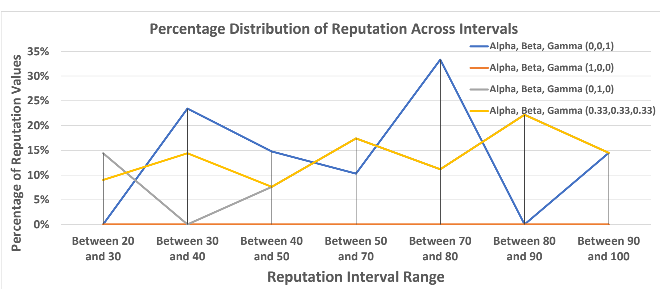

Figure 13: Analysis of Reputation Value Distribution Across Different Parameters in RSMA Blacklisting Framework

D. Interpretation of Modified Reputation Function in RSMA Blacklisting Technique

We conduct a detailed empirical analysis of an enhanced Blacklisting technique by modifying the reputation function used in Reputation System-Based Lightweight Message Authentication (RSMA). The reputation function is articulated as follows:

$$\begin{aligned} RS_{\text{new}} = {} & \beta \times \text{ObjectiveEvaluation} \\ & + \gamma \times \text{EffectDegree} \\ & + \alpha \times \text{HistoricalReputation} \end{aligned} \tag{3}$$

The primary objective was to evaluate the effects of Historical Reputation, Objective Evaluation, and Effect Degree on the reputation score, by adjusting their weights. Figure 13 presents the distribution of reputation values under different parameter settings:

1) **Impact of Objective Evaluation Alone**: The parameter setting $(0, 1, 0)$, accentuating Objective Evaluation, consistently yields lower reputation values, underscoring its relatively minimal impact when functioning in isolation.
2) **Impact of Isolated Effect Degree**: The $(0, 0, 1)$ setting, prioritizing Effect Degree, shows variable performance, indicating that when isolated, Effect Degree can lead to fluctuating reputation scores, showcasing its dynamic nature.
3) **Impact of Historical Reputation Alone**: With parameters set to $(1, 0, 0)$, focusing exclusively on Historical Reputation, the results indicate a consistent distribution of reputation values. This uniformity stems from the factor's reflection of accumulated historical data, maintaining its original value even when other components are nullified, thus highlighting its stabilizing effect.
4) **Integration of All Parameters**: The balanced parameter setting $(0.33, 0.33, 0.33)$, integrating all three components equally, results in a well-distributed range of reputation values. This suggests a synergistic effect achieved through a holistic consideration of all factors.

These insights highlight the distinct roles played by Historical Reputation, Objective Evaluation, and Effect Degree in determining reputation scores within the RSMA Blacklisting framework.

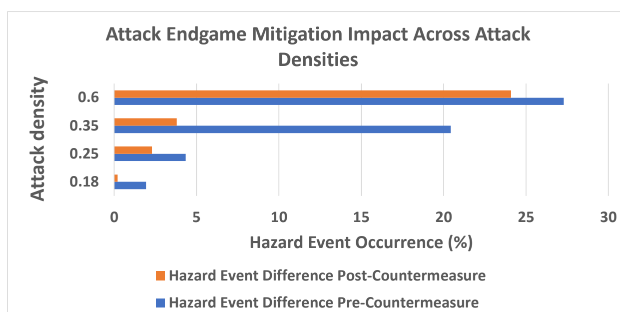

Figure 14: Comparative Impact of Countermeasures on Attack Endgame Across Varying Attack Densities

E. Analyzing Mitigation Efficacy Across Diverse Attack Densities

This experiment was designed to examine the adaptability of our proposed mitigation strategies across varying levels of attack density. The primary objective was to understand the influence of fluctuating attack prevalence on the system's

Table III: Effectiveness of Mitigation Techniques against Fake Accident Attack on Speed Recovery

Mitigation Applied	**Original**	**Blacklisting**	**Geofencing**	**Endgame Prevention**	**ALL**
Message Count Variation (%)	0.0	0.0243	-0.0138	-0.0143	-0.0263
Attack Message Ratio (%)	0.0042	0.0045	0.0043	0.0043	0.0048
Hazard Event Ratio (%)	0.0434	0.0372	0.0366	0.0277	0.0289
Speed Recovery Ratio (%)	1.0	1.2	1.4	2.2	3.0

susceptibility and to evaluate the effectiveness of our countermeasures in reducing the occurrence of hazardous events.

Figure 14 presents a comparative analysis of the impact of countermeasures on the attack endgame within different attack density environments, quantified by the variation in hazard event occurrences. The graph delineates the scenarios before and after the implementation of countermeasures, indicated by light blue and dark blue, respectively. In the pre-countermeasure state, there is a gradual escalation in attack density corresponding to the increase in hazard event occurrences, peaking at approximately 27.28% for the highest level of hazard events noted. Conversely, following the deployment of countermeasures, there is a conspicuous decrease in attack density across all evaluated categories. This is exemplified by the data point at a 5% hazard event occurrence, where the attack density diminishes from 4.34 to 2.28 after countermeasure implementation, signifying a substantial mitigation impact.

The results emphasize the dynamic adaptability of our mitigation strategies, demonstrating their varying degrees of efficacy across different attack densities. Particularly at an attack density of 0.25, it shows the necessity of fine-tuning the countermeasures to enhance their responsiveness and effectiveness in mitigating hazards during phases of transitional attack prevalence.

F. Impact of Mitigation on Message Exchange and Speed Recovery

This section examines how our model affects key operational aspects: the exchange of messages and the recovery of vehicle speeds. These experiments highlight the real-world value and benefits of our mitigation strategies.

Table III shows the findings on the interplay between mitigation techniques on Message Count Variation, Attack Message Ratio, Hazard Event Ratio, and their collective influence on the Speed Recovery Ratio under different mitigation strategies. Speed recovery is defined as the difference between a vehicle's speed profile while a Fake-Reporting Attack is in progress and its speed profile after the proposed mitigation is applied.

1. Message Count Variation: The implementation of Geofencing, Endgame Prevention, and the combined application of all strategies results in a decrease in message count. This reduction, particularly notable in the ALL category with -0.0263%, suggests that a comprehensive mitigation strategy might lead to more efficient message handling, thereby positively impacting speed recovery. Interestingly, the variation in message count shows a slight increase with the application of Blacklisting, indicating a marginal elevation in message traffic which will be considered in our future work.

2. Attack Message Ratio: Across all mitigation methods, there is a slight increase in the attack message ratio, with the combined approach (ALL) showing the highest increase to 0.0048%. This increment might be indicative of the system's heightened sensitivity in detecting attack messages, thereby enhancing the overall robustness of the network and contributing to improved speed recovery.

3. Hazard Event Ratio: A notable decrease in the hazard event ratio is observed with each mitigation strategy, particularly with Endgame Prevention and the combined application of all methods. The decrease to 0.0277% and 0.0289%, respectively, underscores the effectiveness of these strategies in mitigating hazard events, which is a crucial factor leading to an enhanced speed recovery ratio.

4. Speed Recovery Ratio: The speed recovery ratio significantly improves under each mitigation strategy, with the most pronounced improvement observed when all strategies are employed simultaneously (a 200% increase to 3.0%).

In summary, the results suggest that while individual mitigation strategies like Blacklisting, Geofencing, and Endgame Prevention each contribute to improving the speed recovery ratio, their integrated application amplifies this effect. The combined approach enhances the system's efficiency in handling messages and significantly reduces hazard events, thereby enabling a more robust and rapid recovery of vehicle speeds post-attack. All experimental results were repeated 10 times, and the averages were reported. Since the confidence intervals were small, they were not explicitly shown in order to maintain clarity in the presentation of results.

VII. Conclusion and Future Work

This work substantiates the efficacy of a multi-faceted mitigation strategy against cyber-attacks on Vehicle Ad-hoc Networks (VANETs). Our work is novel in targeting attack endgames by predicting the consequences of misbehavior and applying mitigation strategies. Unlike prior research that focuses only on network-level effects, we also consider physical safety outcomes such as congestion or collision risks. This dual perspective ensures that both technical performance and human safety are addressed in VANET security. Leveraging the VeReMi for Attack Prediction (VeReMiAP) dataset and harnessing recent progress in predicting the effects of attacks, our study puts forward a dual-faceted strategy designed to counteract the culmination of attack scenarios effectively. This dual strategy commences with preventing attack endgames, which is achieved through deploying sophisticated Recurrent Neural Network (RNN) models. These models play a critical role in forecasting the potential consequences of False Reporting Attacks generated by fraudulent Cooperative Awareness Messages (CAMs), thus enabling proactive interventions at an early stage. The second phase of our strategy is dedicated to implementing countermeasures for attack endgames. This involves the integration of Geofencing, which utilizes the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm for strategic message regulation. In synergy with this, Dynamic Blacklist Management, anchored in a Reputation System-Based Lightweight Message Authentication (RSMA) framework, is adeptly employed. This system is instrumental in identifying and excluding malevolent vehicles from the network, thereby fortifying the security framework of VANETs against potential threats. The combined use of Endgame Prevention, Blacklisting, and Geofencing reduces road hazards by up to 40%, a clear safety benefit. Together, these techniques shorten abnormal traffic conditions and boost speed recovery, delivering a 200% increase in the rate at which vehicles return to normal driving after an attack.

For future work, we aim to validate our solution in real-world vehicle with the help of existing controlled environments for testing VANET such as the *Area X.O* test track [54] or the *16 Avenue North Test Bed* [55]. We aim to evaluate its effectiveness against unpredictable human behaviors, including non-compliance with traffic laws, erratic driving, varied reactions to system alerts, differences in technology adoption, and psychological responses to warnings. By exploring these factors, we seek to refine our system's real-world applicability. Further, We plan to expand our solution to include the Recovery phase of attack mitigation, aiming to develop complete system restoration methods after an attack. Additionally, we intend to broaden our analytical scope to cover a wider range of adversarial tactics, including the Sybil attack, a threat wherein attackers create multiple fake identities to compromise network integrity. We also plan to investigate scenarios involving the control of emergency vehicles through fraudulent CAMs.

REFERENCES

[1] M. A. Al-Shareeda and S. Manickam, "A systematic literature review on security of vehicular ad-hoc network (VANET) based on veins framework," *IEEE Access*, pp. 46 218–46 228, 2023.

[2] G. Dimitrakopoulos and P. Demestichas, "Intelligent transportation systems," *IEEE Vehicular Technology Magazine*, pp. 77–84, 2010.

[3] Y. Cao, O. Kaiwartya *et al.*, "Guest editorial special issue on toward securing internet of connected vehicles (iov) from virtual vehicle hijacking," *IEEE Internet of Things Journal*, pp. 5866–5869, 2019.

[4] A. Gaurav, B. B. Gupta *et al.*, "A novel approach for fake news detection in vehicular ad-hoc network (VANET)," in *Computational Data and Social Networks*. Springer International Publishing, 2020, pp. 386–397.

[5] M. A. Abdelmaguid, H. S. Hassanein, and M. Zulkernine, "Attack endgame: Proactive security approach for predicting attack consequences in VANET," in *ICC IEEE International Conference on Communications*, 2023, pp. 3762–3767.

[6] A. K. Malhi, S. Batra, and H. S. Pannu, "Security of vehicular ad-hoc networks: A comprehensive survey," *Computers & Security*, p. 101664, 2020.

[7] Y. Hao, Y. Cheng *et al.*, "A distributed key management framework with cooperative message authentication in vanets," *IEEE Journal on selected areas in communications*, pp. 616–629, 2011.

[8] H. Amari, Z. A. E. Houda *et al.*, "Trust management in vehicular ad-hoc networks: Extensive survey," *IEEE Access*, pp. 47 659–47 680, 2023.

[9] Y. Liang, E. Luo, and Y. Liu, "Physically secure and conditional-privacy authenticated key agreement for VANETs," *IEEE Transactions on Vehicular Technology*, pp. 7914–7925, 2023.

[10] Y. Zhuang, J. Pan *et al.*, "Time and location-critical emergency message dissemination for vehicular ad-hoc networks," *IEEE Journal on Selected Areas in Communications*, pp. 187–196, 2010.

[11] S. Biswas, M. M. Haque, and J. V. Misic, "Privacy and anonymity in vanets: A contemporary study." *Ad Hoc Sens. Wirel. Networks*, pp. 177–192, 2010.

[12] B. K. Mekonen, L. Bane, and N. B. Fite, "Detection of false position attacks in VANETs through bagging ensemble learning," *PLoS One*, p. e0328829, 2025.

[13] S. Gyawali, Y. Qian, and R. Q. Hu, "A privacy-preserving misbehavior detection system in vehicular communication networks," *IEEE Transactions on Vehicular Technology*, pp. 6147–6158, 2021.

[14] A. Sharma and A. Jaekel, "Machine learning based misbehaviour detection in VANET using consecutive bsm approach," *IEEE Open Journal of Vehicular Technology*, pp. 1–14, 2021.

[15] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, p. 132306, 2020.

[16] C. Kumar and S. Mohanty, "Current trends in cyber security for drones," in *International Carnahan Conference on Security Technology (ICCST)*, 2021, pp. 1–5.

[17] Q. Liu, S. Yan *et al.*, "Multi-malicious nodes double-spending attack blacklist management model," in *4th International Conference on Data Science and Information Technology*, 2021, pp. 462–467.

[18] G. Dong, Y. Jin *et al.*, "Db-kmeans: an intrusion detection algorithm based on dbscan and k-means," in *20th Asia-Pacific Network Operations and Management Symposium (APNOMS)*. IEEE, 2019, pp. 1–4.

[19] J. Cui, X. Zhang *et al.*, "Rsma: Reputation system-based lightweight message authentication framework and protocol for 5g-enabled vehicular networks," *IEEE Internet of Things Journal*, pp. 6417–6428, 2019.

[20] S. Barghi, A. Benslimane, and C. Assi, "A lifetime-based routing protocol for connecting vanets to the internet," in *2009 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops*, 2009, pp. 1–9.

[21] S. Dasgupta, M. Rahman *et al.*, "Prediction-based GNSS spoofing attack detection for autonomous vehicles," *arXiv preprint arXiv:2010.11722*, 2020.

[22] F. Laghrissi, S. Douzi *et al.*, "Intrusion detection systems using long short-term memory (lstm)," *Journal of Big Data*, p. 65, 2021.

[23] H. Schafer, E. Santana *et al.*, "A commute in data: The comma2k19 dataset," *arXiv arXiv:1812.05752*, 2018.

[24] P. T. Yamak, L. Yujian, and P. K. Gadosey, "A comparison between arima, lstm, and gru for time series forecasting," in *Proceedings of the 2nd International Conference on Algorithms, Computing and Artificial Intelligence*, 2019, pp. 49–55.

[25] G. D. Putra and S. Sulistyo, "Trust based approach in adjacent vehicles to mitigate sybil attacks in VANET," in *Proceedings of the International Conference on Software and E-Business*, 2017, p. 117–122.

[26] M. Tsukada, S. Arii *et al.*, "Misbehavior detection using collective perception under privacy considerations," in *IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, 2022, p. 808–814.

[27] J. Kang, Z. Xiong *et al.*, "Toward secure blockchain-enabled internet of vehicles: Optimizing consensus management using reputation and contract theory," *IEEE Transactions on Vehicular Technology*, pp. 2906–2920, 2019.

[28] L. M. Santos-Jaimes and E. d. S. Moreira, "Pseudonym change strategy based on the reputation of the neighbouring vehicles in vanets," *Dyna*, pp. 157–166, 2019.

[29] X. Yang, F. Zhu *et al.*, "Secure reputation-based authentication with malicious detection in VANETs," *IEEE Transactions on Dependable and Secure Computing*, pp. 359–372, 2024.

[30] R. Hussain, J. Lee, and S. Zeadally, "Trust in VANET: A survey of current solutions and future research opportunities," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 2553–2571, 2020.

[31] A. Mahmood, Q. Z. Sheng *et al.*, "Toward a distributed trust management system for misbehavior detection in the internet of vehicles," *ACM Trans. Cyber-Phys. Syst.*, 2023.

[32] T. Alladi, B. Gera *et al.*, "Deepadv: A deep neural network framework for anomaly detection in vanets," *IEEE Transactions on Vehicular Technology*, pp. 12 013–12 023, 2021.

[33] M. Maad Hamdi, L. Audah *et al.*, "A review of applications, characteristics and challenges in vehicular ad hoc networks (VANETs)," in *2020*

International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2020, pp. 1–7.

[34] M. Husák, J. Komárková *et al.*, "Survey of attack projection, prediction, and forecasting in cyber security," *IEEE Communications Surveys & Tutorials*, pp. 640–660, 2018.

[35] J. Yang, Y. Guo *et al.*, "Vulnerability analysis of road network under information pollution attacks in vanet," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[36] S. A. Asra, "Security issues of vehicular ad hoc networks (VANET): A systematic review," *TIERS Information Technology Journal*, pp. 17–27, 2022.

[37] G. A. Jagnade, S. I. Saudagar, and S. A. Chorey, "Secure VANET from vampire attack using LEACH protocol," in *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, 2016, pp. 2001–2005.

[38] I. Elleuch, A. Makni, and R. Bouaziz, "Towards a cooperative intelligent system for unpredictable and predictable road hazard detection," *Int. J. Comput. Inf. Syst. Ind. Manag. Appl*, pp. 101–114, 2019.

[39] I. Elleuch, A. Makni, and R. Bouaziz, "Design of an intelligent cooperative road hazard detection persistent system," in *Intelligent Systems Design and Applications*, 2020, pp. 696–707.

[40] S. Shivshankar and A. Jamalipour, "An evolutionary game theory-based approach to cooperation in vanets under different network conditions," *IEEE Transactions on Vehicular Technology*, pp. 2015–2022, 2014.

[41] R. Sedar, C. Kalalas *et al.*, "Knowledge transfer for collaborative misbehavior detection in untrusted vehicular environments," *IEEE Transactions on Vehicular Technology*, 2024.

[42] M. H. Shahriar, M. R. Ansari *et al.*, "VehiGAN: Generative Adversarial Networks for Adversarially Robust V2X Misbehavior Detection Systems," *ACM Transactions on Cyber-Physical Systems*, pp. 1–26, 2025.

[43] T. Anjali, R. Goyal, and G. Balaji, "X-GEVON-a novel explainable intelligent network to detect the multiple attacks in VANET systems," *Discover Computing*, p. 185, 2025.

[44] X. Xu, Y. Wang, and P. Wang, "Comprehensive review on misbehavior detection for vehicular ad hoc networks," *Journal of Advanced Transportation*, p. 4725805, 2022.

[45] M. A. Abdelmaguid, H. S. Hassanein, and M. Zulkernine, "A VeReMi-based dataset for predicting the effect of attacks in VANETs," in *Proceedings of the Int'l ACM Conference on Modeling Analysis and Simulation of Wireless and Mobile Systems*, 2023, p. 175–179.

[46] M. A. Abdelmaguid, H. S. Hassanein, and M. Zulkernine, "VeReMiAP: A VeReMi-based Dataset for Predicting the Effect of Attacks in VANETs," 2023. [Online]. Available: https://doi.org/10.5683/SP3/R09EWA

[47] L. Codecà, R. Frank, and T. Engel, "LuST: a 24-hour scenario of Luxembourg city for SUMO traffic simulations," in *SUMO User Conference - Intermodal Simulation for Intermodal Transport*, 2015.

[48] R. Sedar, C. Kalalas *et al.*, "Misbehavior detection in vehicular networks: An ensemble learning approach," in *IEEE Global Communications Conference*. IEEE, 2022, pp. 1850–1855.

[49] H. Li, K. Kalogiannis *et al.*, "AttentionGuard: Transformer-based Misbehavior Detection for Secure Vehicular Platoons," *arXiv preprint arXiv:2505.10273*, 2025.

[50] Z. Xu, X. Li *et al.*, "Dsrc versus 4g-lte for connected vehicle applications: A study on field experiments of vehicular communication performance," *Journal of Advanced Transportation*, 2017.

[51] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved IVC analysis," *IEEE Transactions on Mobile Computing*, pp. 3–15, 2011.

[52] D. Krajzewicz, J. Erdmann *et al.*, "Recent development and applications of sumo-simulation of urban mobility," *International Journal on Advances in Systems and Measurements*, pp. 128–138, 2012.

[53] Al-ani, Ruqayah and Zhou, Bo and Shi, Qi and Baker, Thar and Abdlhamed, Mohamed, "Adjusted Location Privacy Scheme for VANET Safety Applications," in *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*, 2020, pp. 1–4.

[54] "Area x.o smart mobility test facility," https://areaxo.com/, Area X.O, 2025, accessed 26 Apr 2025.

[55] "16 avenue north connected vehicle test bed," https://science-ouverte.canada.ca/items/8b111cb2-53f8-42e3-8b76-dc84fb17626d, Transport Canada, 2023, accessed 26 Apr 2025.