IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 June 2023, accepted 26 June 2023, date of publication 4 July 2023, date of current version 11 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3292055

RESEARCH ARTICLE

A Novel Mechanism for Misbehavior Detection in Vehicular Networks

EDIVALDO PASTORI VALENTINI[1], GERALDO PEREIRA ROCHA FILHO[2], ROBSON EDUARDO DE GRANDE[3], (Member, IEEE), CAETANO MAZZONI RANIERI[4], LOURENÇO ALVES PEREIRA JÚNIOR[5], AND RODOLFO IPOLITO MENEGUETTE[4], (Member, IEEE)

[1]Federal Institute of Education, Science and Technology of São Paulo, IFSP, Catanduva, São Paulo 15808-305, Brazil
[2]Department of Exact and Technological Sciences, State University of Southwest Bahia (UESB), Vitória da Conquista, Bahia 45083-900, Brazil
[3]Department of Computer Science, Brock University, St. Catharines, ON L2S 3A1, Canada
[4]Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, São Paulo 13566-590, Brazil
[5]Department of Computer Science, Technological Institute of Aeronautics, São Jose ́ dos Campos, São Paulo 12228-900, Brazil

Corresponding author: Rodolfo Ipolito Meneguette (meneguette@icmc.usp.br)

The authors would like to thank the CEPID-CeMEAI/ICMC-USP, the support from the São Paulo Research Foundation (FAPESP) grant #2020/07162-0, #2021/06210-3, #2021/10921-2, #2013/07375-0, #2023/08420-0) and the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery grant).

ABSTRACT Intelligent Transport Systems (ITS) have provided new technologies to protect human life, speed up assistance, and improve traffic, to aid drivers, passengers, and pedestrians. Vehicular Ad-hoc Networks (VANET) are the fundamental elements in an ITS ecosystem. However, its characteristics make the system susceptible to numerous attacks, such as Denial of Service (DoS). In this paper, we proposed a security system based on intrusion detection called Detection of Anomalous Behaviour in Smart Conveyance Operations (DAMASCO). We used a statistical approach to detect anomalies in vehicle-to-vehicle communication (V2V). The anomaly detection module addresses the Medium Access Control (MAC) sublayer to assess the number of packages sent to identify potentially malicious nodes, block their activity, and maintain a reputation list. The algorithm calculates the Median Absolute Deviation (MAD) to identify outliers and characteristics of DoS. Our experiments were performed in a simulated environment using a realistic urban mobility model. The results demonstrate that the proposed security system achieved a 3% false positive rate and no false negatives.

INDEX TERMS Cyber attacks, denial of service, Internet of Things, intrusion detection, smart cities, vehicular networks.

I. INTRODUCTION

The Intelligent Transport System (ITS) paradigm plays a crucial role in technological advances and innovations in transport [30], [41]. This paradigm gradually establishes a cooperative and collaborative ecosystem favoured by vehicles' information exchange, road communication infrastructures, cloud services, and internet connections. Among the benefits provided by ITS are tools for traffic jam management and traffic efficiency [30], [40], [51].

Vehicles are the fundamental elements in an ITS ecosystem, along with the Road Side Units (RSU), responsible for the communication infrastructure installed throughout the highways. The so-called Vehicular Ad-Hoc Networks (VANET) enable a distributed and cooperative infrastructure for data exchange [14], [19], [58]. Communication between VANET nodes can be established following approaches: Vehicles to Vehicles (V2V), Vehicles to Road Infrastructure (V2I), and Vehicles to Everything (V2X) [42], [50].

Malicious vehicles or intruders can carry out deviant behaviour, damaging the transport system in different ways: (i) causing data congestion on the network, (ii) sending false alerts, (iii) performing identity theft, (iv) spoofing

The associate editor coordinating the review of this manuscript and approving it for publication was Xianzhi Wang.

location, and (v) exploiting vulnerabilities in the routing protocol [8], [24], [27], [29], [31], [53]. Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks can inflict critical consequences in real-time applications, causing delays in ITS routines such as emergency and rescue operations [15], [18], [61].

ITS applications often have limited security requirements, generating the risk for the computing infrastructure to host malware that can harmfully exploit devices [21], [29]. An example is Mirai, a malware that can be manipulated remotely through a network of malicious devices and systems, referred to as botnets [7], [32]. These botnets can initiate thousands of DoS/DDoS attacks on any susceptible device [13], [17], [47].

There are works in the literature [1], [11], [22], [25], [28], [37], [38], [39], [49], [59], [64] that deal with the identification of anomalies in the network. However, these works do not consider the identification of anomalies that occur due to the misbehaviour of the data link. Furthermore, some of these works do not use any list to indicate whether the vehicle is a possible malicious vehicle.

Our work addressed this research gap by proposing a detection and mitigation method. We introduced a security system called Detection of Anomalous Behaviour in Smart Conveyance Operations (DAMASCO). This system identifies attacks during information exchanges between vehicles and the VANET infrastructure. Utilising an anomaly detection technique, DAMASCO identifies harmful behaviour originating from the data link layer of malicious vehicles. The system uncovers anomalies tied to DoS and DDoS attacks using a statistical model that identifies extreme values in a decentralised setting. Simulation results revealed the effectiveness of DAMASCO, demonstrating a high detection rate of 95%, with a low false positive rate of 3% and no false negatives.

The main contributions of this work are twofold:

- Design and evaluation of a decentralised and cooperative security system for detecting DoS/DDoS attacks in VANET networks based on Medium Access Control (MAC) frames and reputation lists, considering unique characteristics of a DoS/DDoS attacks such as transmission storm and flood;
- Design and evaluation of an anomaly detection module based on a straightforward statistics technique applied to broadcast requests.

The remainder of the article is structured as follows. Section II discusses related works related to this research topic. Section III presents an overview of DAMASCO. Section IV describes the methodology used to evaluate DAMASCO. Finally, Section V presents conclusions and directions for future works.

II. RELATED WORK

An Intrusion Detection System (IDS) can be composed of hardware, software, or a combination of both. Its primary function is to monitor activities within computer networks or systems, scanning for unauthorised operations or violations of information security policies [60]. To categorise various types of threats, attacks, and aberrant activities in traditional networks, systems, or VANETs, an IDS may employ several detection techniques. These include anomaly-based, signature-based, watchdog, or cross-layer techniques. An IDS may also use a combination of these techniques, resulting in what is known as a hybrid IDS [16], [54], [57].

As already stated, vehicular networks are challenging because the vehicles operate in a completely mobile environment, their connection intervals are small, and, in most cases, communication is wireless [24], [31], [53]. Furthermore, their topology is highly dynamic. The testing processes and the prototype development are non-trivial due to the lack of similar or compatible scenarios with real traffic and transport scenarios [57].

Assessing the conditions of the IDS implementation and execution location, particularly in vehicles and RSUs, constitutes a crucial phase in the development of the security system and the selection of algorithms for attack detection and classification. Models and statistical methods for identifying and categorising threats encompass (i) indicators of compromise arising from events generated by systems, hosts, and networks [54], [57], (ii) parametric and non-parametric techniques, (iii) computation of statistical metrics such as mean and variance, (iv) approximation of Bayesian networks via probabilistic relationships, and (v) detection of outliers [33], [44]. An IDS typically undergo validation through computer simulation before actual deployment [53], [54]. In terms of performance evaluation, the most frequently reported metrics are the rates of false positives and false negatives [4], [9].

Lyamin et al. [39] focused on real-time detection of attacks in VANETs. The method they proposed evaluated periodic position messages (beacons) shared between vehicles in a platoon of up to 25 nodes. The authors assessed Jammer-type attackers, who interfere with or corrupt data packets transmitted by the VANET. Two kinds of attacks were considered: packet corruption in the transmission channel and beacon corruption. Utilising statistical models that incorporated mean values, recursive functions, and probabilities, they achieved a detection rate of 90% with no false positives. However, a limitation of this method was its scope, as it only analysed up to 25 nodes per platoon. As a result, it failed to identify, for instance, a DDoS attack occurring outside the platoon.

In their study, Zaidi et al. [64] proposed a system that detected and categorised unusual activities and rogue nodes in vehicle networks. The method aimed to identify false information, including Sybil attacks, in which a malicious entity impersonates multiple false vehicles. To do this, vehicles collected data from nearby vehicles and analysed road traffic density and flow rate. Hypothesis testing was used to verify the authenticity of the data. They used various software tools for network simulations, traffic generation, and car mobility management. They tested their method in four scenarios with varying levels of malicious vehicle presence, achieving a high detection rate of 97.5% with only a 2.5% false positive rate. However, their method only addressed false information and Sybil attacks.

In Loukas et al. [37], an IDS designed for a robotic vehicle was proposed. Such a scenario poses limitations in storage and processing capabilities, resulting in increased latency during the detection process. The authors developed a model based on binary classification, where data tuples were classified as either attacks or non-attacks. By identifying the minimum number of data points required for the detection process, the study focused on the more intricate command injection attack. The findings revealed that offloading certain operations was beneficial. However, there was a notable overall latency due to data collection, network transmission, classification and response, along with factors such as vehicle type, environment, and other characteristics. The authors also highlighted the limitations imposed by machine learning techniques in terms of critical time.

In their study, So et al. [59] focused on the detection of inadequate compartmentalisation in vehicles by analysing the Received Signal Strength Indicator (RSSI). The research primarily dealt with operations pertaining to the physical layer and employed a plausibility model to identify instances of falsified node locations. Safety messages are transmitted by vehicles to convey their respective locations. The detection process relied on V2V communication, where each vehicle independently performed the analysis. The researchers enhanced an existing dataset named the Vehicular Reference Misbehaviour Dataset (VeReMi), which comprised five types of attacks, three vehicle densities, and variations in the attacking vehicles. The study achieved a detection rate of 93% for the Weighted-BSM model, with a recurrence rate of 83%. The First-BSM model demonstrated the highest recurrence rate, reaching approximately 84%.

The work of Hamdan et al. [22] focused on the detection of Sybil attacks, devising a hybrid detection methodology. This approach amalgamated the detection of pseudonymous abuse, which demonstrates increased effectiveness as vehicle count rises, and the privacy-based footprinting algorithm, noted for its superior performance with escalating vehicle speed. However, this hybrid fusion's capacity to discern false identities necessitates cryptographic security measures. These include node authentication, cryptographic hash computation, and the capture of vehicle identifications and their trajectories. These operations, while enhancing the detection rate, can introduce latency and impinge upon the performance of IDS execution and monitoring due to the need for cryptographic calculations and hash analysis. The IDS relies on RSUs, stationed along the route, to process vehicle authentication requests, carry out verification, and perform cryptographic processing at specific intervals.

Luong et al. [38] proposed a security routing protocol to detect and prevent flood attacks. Their approach utilised the median filter statistical method and was anchored in the workings of the Ad-hoc On-Demand Distance Vector (AODV) routing protocol. The security system hinged on three factors for flood detection: (i) monitoring of route discovery to assess the node's historical behaviour; (ii) implementation of the non-linear mapping via the median filter to create a novel analysis space; and (iii) application of robust statistics using the median value in this distinct data sample space for the classification of malicious and normal nodes. The detection of malicious nodes deployed the median value of the route discovery frequency vector, comparing it to a fixed threshold to ascertain whether the node is normal or malicious. Simulation experiments, conducted using Network Simulator 2 (NS-2) and V2V communication, showed the algorithm to be efficacious in detecting flood attacks. However, these simulations were limited to a number of malicious nodes varying from two to six vehicles. The authors did not address historical records within vehicles or reputation lists for future queries, which could potentially reduce the time required for behavioural and historical analysis of neighbouring vehicles.

Paranjothi and Atiquzzaman [49] designed a statistical method to detect rogue nodes transmitting false data, leveraging the principles of FOG computing. A central vehicle termed a guard node was selected to operate as a FOG layer, processing messages from other vehicles' onboard units (OBUs). The data collected, such as velocity and position, was processed using statistical methods for rogue vehicle detection. The approach was tested using OMNET+ and SUMO simulators across two mobility scenarios. The method didn't account for DoS or flood attacks. Besides, it was heavily reliant on guard nodes for maintaining the IDS and communicating detected rogue nodes.

Haydari and Yilmaz [25] presented a security system for vehicular networks addressing false data injection and DDoS attacks. Their system, an IDS for anomaly detection, integrated machine learning (k-nearest neighbours - kNN) and statistical methods (cumulative sum control - CusUM). This processing occurred exclusively at Road Side Units (RSUs). The system was developed using simulators, with a focus on VANET communication and realistic traffic scenarios. However, simulations were restricted to straight two-way traffic sections, with no consideration for complex urban road networks. While encrypted messages were assumed, the paper didn't specify the encryption type used. The system lacked reputation or voting lists for trusted node election, solely blocking transmissions from identified malicious vehicles. As the security was fully centralised in roadside infrastructure, vehicles relied on RSU-processed decisions to determine the trustworthiness of neighbouring nodes.

Vinita and Vetriselvi [28] proposed the Federated Learning Entrusted Misbehaviour Detection System (FLEMDS) with vehicle selection to support the 6G-enabled vehicles in combating Sybil attackers. As a result, Sybil attack detection is carried out locally in the vehicles employing the federated learning on-vehicle AI technique. The FLEMDS employs a three-level model weight aggregation process at three locations to improve detection accuracy. To minimise the learning and detection latency, federated learning and software-defined vehicular fog computing are combined in the FLEMDS. For this aim, the approach is based on a fuzzy logic-based FL-vehicle selection technique to choose suitable vehicles as clients for the participation of local training in the FL process.

Chakraborty et. al [11] proposed a security technique that improves communication and intruder detection in VANET for smart transportation. Ciphertext-policy game theory encryption analysis for smart transportation is used here to improve the security of the VANET. Fuzzy rule-based encoder perceptron neural networks are utilised in the detection of the VANET intruder. However, this approach brings a high conversation time given the combination of fuzzy and neural networks.

Shams et. al [1] proposed a multi-class intrusion detection system using Convolutional Neural Network with a novel feature extraction method known as Context-Aware Feature Extraction-Based CNN (CAFECNN). The CAFECNN model takes advantage of the collected network flow data to detect both passive and active types of attacks.

The related works in this section considered different architectures and detection techniques, from RSU-base to fully distributed, capable of handling one or more attacks. Table 1 summarises the most prominent aspects of these works and highlights the contributions of our security model. Our criteria encompass deployment locations, detection techniques, threat classification models, and other characteristics.

In Lyamin et al. [39], the authors addressed DoS detection as a corruption of beacons and data from the transmission channels and did not specify DDoS attacks and flood attacks. Zaidi et al. [64] and Paranjothi and Atiquzzaman [49] did not implement reputation processes nor addressed DDoS attacks (the detection process depended on sending parameters of the location and speed of neighbouring vehicles). So et al. [59] was specific to the physical layer, analysing the RSSI power. To be more precise, the detection technique depended on information shared by the user. However, this work did not explore DoS or DDoS attacks. Hamdan et al. [22] only detected the Sybil attack using false identities. A hybrid detection method combines cryptographic methods and vehicle privacy analysis. However, these executions increase the latency of the detection rate and overload the IDS. RSU support needs to process vehicle authentication requests and verification and generate encryption.

In Paranjothi and Atiquzzaman [49], the method is highly dependent on the central guardian nodes for the IDS to run and broadcast which vehicles have been detected as rogues. Differently, Loukas et al. [37] and Haydari and Yilmaz [25] used a centralised architecture to detect command injection, false data, and DDoS attacks. Furthermore, these approaches used machine learning techniques with a binary classification process [37], kNN algorithms and statistical methods that can insert some latency between the vehicle security application and the remote storage architecture or centralised. The IDS of Luong et al. [38] updated the AODV protocol by implementing the statistical median filter method at the network layer. In the simulation, they applied only 2 to 6 malicious vehicles. The work does not describe where the history of monitored vehicles is stored, i.e., it does not implement a reputation list for future consultations.

These related works do not consider the identification of anomalies that occur due to the misbehaviour of the data link. Furthermore, the works in the literature do not use any list to indicate whether the vehicle is a possible malicious vehicle. Therefore, this paper proposed a security solution that performs anomaly detection by analysing the data link layer packet transmission pattern through a statistical model. For this, the proposed solution monitors the link layer messages transmitted in each. If the transmission deviates from the request-sending pattern due to an attempt of a DoS/DDoS attack, and if it is identified, the transmission is classified. and the vehicle that caused the transmission is identified as a malicious vehicle.

Our method provides a significant contribution by offering a faster approach to classifying normal and abnormal network traffic, with reduced computational processing. This type of analysis was previously conducted in only one study from the existing literature [22]. However, the said study envisaged a centralised situation where all packets would pass through a single infrastructure. This is different from a vehicular infrastructure, where, by default, numerous packets pass through each vehicle. These packets do not necessarily represent an attack but rather constitute information from the vehicular network.

III. DAMASCO: DETECTION OF ANOMALOUS BEHAVIOUR IN SMART CONVEYANCE OPERATIONS

The Detection of Anomalous Behaviour in Smart Conveyance Operations within Vehicular Networks (DAMASCO) consists of an intrusion detection structure for VANETs. The following subsections describe the system development, present the statistical model for anomaly detection, and describe the processes for monitoring, analysing, and collecting data.

A. SYSTEM OVERVIEW

DAMASCO aims to identify DoS attacks in a VANET, minimising the risks of accidents caused by malicious actions by attackers. Its operations aim to protect drivers, passengers, and pedestrians in urban environments, where the density of vehicles is high.

The presented system exclusively focused on V2V communication, eschewing any supplementary infrastructure or external aid for communication or transmission handling. Furthermore, it refrained from creating vehicular clusters, thus enabling each vehicle to individually assess traffic patterns and categorise proximate vehicles. This scenario allowed for more rapid and dynamic anomaly detection. It also reduced the transmission time to a peripheral base

TABLE 1. IDS-based security solutions for VANETs and ITS.

		Detection methodologies		Main goals and Features	
Work	Deployment	Anomaly	Classification	MAC Layer	Reputation List
Lyamin et al. [39]	Distributed	✓	Statistics		
Zaidi et al. [64]	Distributed	✓	Statistics		
Loukas et al. [37]	Centralised	✓	Deep learning		
Vinita and Vetriselvi [28]	Centralised	✓	Fuzzy		
Chakraborty et al. [11]	Centralised	✓	Neural networks		
Shams et al. [1]	Centralised	✓	CNN		
So et al. [59]	Both	✓	Plausibility		
Hamdan et al. [22]	Centralized	✓	Encryption and Authentication	✓	
Luong et al. [38]	Distributed	✓	Statistics		
Paranjothi and Atiquzzaman [49]	Centralized	✓	Statistics		
Haydari and Yilmaz [25]	Centralized	✓	kNN / Statistics		
Ours	Distributed	✓	Outliers	✓	✓

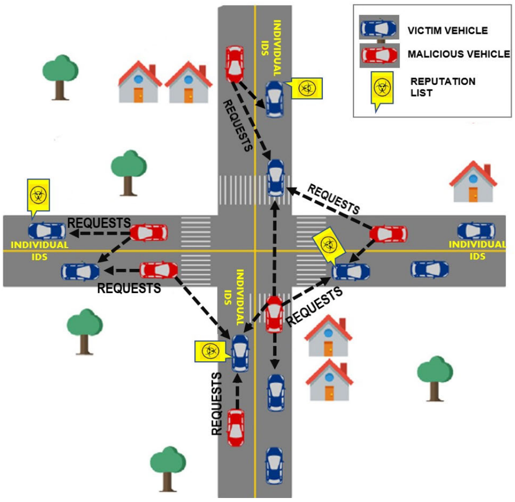

FIGURE 1. Abstraction of the DAMASCO security proposal.

station and hastened the response to potential threats on the vehicular network.

Vehicles establish ad-hoc connections and send requests to each other where malicious vehicles may perform attacks on victim vehicles. Each vehicle can send REQUEST messages to other vehicles within its limits. Malicious vehicles, in red (Figure 1), can carry out the DoS/DDoS attacks on target vehicles in blue. The individual IDS performs anomaly detection to identify these attacks. Upon malicious action detection by any vehicle, the security system stores information regarding the attacking vehicle in a data structure called a reputation list. The reputation list allows for sharing in future interactions so that the other vehicle can consult it and identify whether the neighbouring vehicle is malicious.

DAMASCO is slated for installation and operation within individual vehicles, functioning via a decentralised model and identifying local, or host-based, attacks. As previously depicted, two types of nodes are defined: the victim vehicles and the malicious vehicles. These vehicles are dispersed throughout the urban environment. Regardless of the communication between Road Side Units (RSUs) or Cluster Head (CH) vehicles, each runs the DAMASCO security system autonomously. Wireless communication, conforming to the IEEE 802.11p standard, spans a geographical range of approximately 1 *km*.

B. METHODS FOR ANOMALY DETECTION

Legitimate data is typically drawn from a specific, application-dependent distribution. When the data generation process behaves unusually, abnormal data are likely to produce outliers. Therefore, outliers may be a helpful indication of unexpected behaviour within a system [2]. Our approach aimed at identifying and analysing outliers for the anomaly detection and classification process.

Anomaly detection algorithms can be based on statistics or machine learning [3]. Statistics-based methods work by computing statistical properties in a dataset and, based on a set of assumptions regarding the data distribution, establishing rejection criteria to identify anomalous samples. Machine learning-based methods employ a supervised or unsupervised learning architecture to a set of training samples, inducing a model that can address an unknown distribution and deploying the resulting model.

Machine learning models can handle complex problems, including unstructured data such as images or text [48], besides being able to address structured data with more accurate results than statistical methods [46]. However, they require a costly training phase and must be continuously maintained to provide consistent predictions [55]. Statistical methods are easier to maintain and require hardware/software with minimal performance and processing resources.

Given the volatile dynamics inherent in VANETs data transmission, the incorporation of machine learning models into this class of network would necessitate an in-depth evaluation of each utilised algorithm. Moreover, maintaining the model up-to-date would potentially require constant retraining within a distributed system. Factoring in privacy stipulations, such a pipeline might involve the deployment of federated learning methodologies [35], an aspect that lies outside the purview of this research. Our primary objective was to develop a solution that effectively mitigates

DoS/DDoS attacks predicated on MAC frames and reputation lists, ensuring minimal computational overhead and eliminating the need for centralisation. Consequently, we implemented a simple, yet effective statistical approach capable of autonomous operation within each node: the Median Absolute Deviation (MAD) method [23].

1) THE MAD METHOD

Our approach for outliers' detection uses the MAD method [23], [34], which can detect outliers from any unimodal symmetric distribution [10]. In the case of asymmetric distributions, alternative methods [52] may be introduced to the system without invalidating the overall architecture. Moreover, recent related work has employed methods that assume a symmetric distribution [5], [6] for outlier detection in VANETs to prevent DDOS attacks.

The MAD method considers the median of a set, which is the measure of central tendency most robust to the presence of outliers. The objective is to establish a rejection criterion, characterised as a maximum deviation around the median beyond which the samples might be considered outliers. These rejection criteria are derived from the MAD value, defined in Equation 1, where $\mathbf{x}$ is an input distribution, b is a coefficient, and function M computes the median of a dataset.

$$MAD = b \cdot M(\left\| \mathbf{x} - M(\mathbf{x}) \right\|). \tag{1}$$

This equation defines that the MAD value of a dataset $\mathbf{x}$ is computed by (i) calculating the dataset median $M(\mathbf{x})$, (ii) subtracting it from all elements of $\mathbf{x}$, (iii) taking the absolute values, (iv) calculating the median of the resulting set, and (v) multiplying it by a constant b. Assuming a probability distribution for the data, the method states that $b = \frac{1}{Q(0.75)}$, where $Q(0.75)$ is the 0.75 quantile of this distribution. In the case of a normal distribution, this computation results in $b = 1.4826$ [26].

The subset of normal samples, denoted as $\hat{\mathbf{x}}$, is established as indicated in Equation 2. Here, ec represents an arbitrary constant referred to as the exclusion criterion. The criterion becomes more conservative as its value increases, typically falling within the range $ec \in 2.0, 2.5, 3.0$ [26]. Any elements in $\mathbf{x}$ that are absent in $\hat{\mathbf{x}}$ are categorised as outliers.

$$\hat{\mathbf{x}} \quad \in \quad [M(\mathbf{x}) - ec \cdot MAD, \quad M(\mathbf{x}) + ec \cdot MAD] \tag{2}$$

Within the context of VANETs, the attack model simulated in our experiments presumed the target vehicles were inundated with a substantial number of REQUEST packets. This volume of REQUEST packets was the key parameter for anomaly detection within our study's framework. Hence, outliers demonstrating a positive deviation (i.e., values surpassing the threshold of $M(\mathbf{x}) + ec \cdot MAD$) were treated as signals of anomalous network communication activity.

C. DAMASCO OPERATION

Algorithm 1 summarises the overall reputation analysis in DAMASCO. REQUEST messages received from neighbouring vehicles are inspected, and the metadata of these operations is stored. Periodically, following a predefined interval defined as ϕ, the stored data are analysed to identify possible anomalies inherent to DoS/DDoS. The total number of messages is fed to the MAD method for outlier detection. If patterns indicative of DoS/DDoS attacks are detected, the source vehicle is labelled as malicious and added to the reputation list accordingly. Otherwise, the activity is considered normal behaviour.

Algorithm 1 Overall DAMASCO Operation

```
Require: $ec, \phi, t$
$t \leftarrow init_t()$
$MAC_v \leftarrow init_{MV}()$
$list_{rep} \leftarrow init_l()$
while not $SHUT\_DOWN$ do
    if $LISTEN(LLC_f, MAC_f)$ then
        $proc\_frame(LLC_f, MAC_f, MAC_v)$
    end if
    if $t \ \% \ \phi == 0$ then
        $list_{rep} \leftarrow UPDATE\_LIST(MAC_v, ec)$
        $MAC_v \leftarrow init_{MV}()$
    end if
end while
```

MAC tables underwent scrutiny and analysis through protocol identification within the LLC sublayer, submission of REQUEST frames, and tracking of these frames received by the vehicles. These operations correspond to the following two phases of the system execution cycle:

1) Monitoring and constantly analysing Medium Access Control (MAC) frames received from vehicles within their radio range (Algorithm 2).
    a) When identifying a REQUEST from the MAC frame, the algorithm distinguishes whether the frame carries a REQUEST or a REPLY message;
    b) If it is a REQUEST, then the timestamp, the MAC address and the IPv4 address of the vehicle are collected, and the $total_{req}$ counter is incremented.
2) Performing anomaly detection through the MAD method (Algorithm 3).
    a) The algorithm analyses the total number of REQUEST messages received from each neighbouring vehicle, seeking to identify values considered extreme;
    b) Vehicles that have an abnormal number of packages sent are classified as malicious, blocked, and registered in a local reputation list.

D. DETECTION TECHNIQUE

As already discussed, the detection technique identifies DoS/DDoS attacks by analysing the behaviour of the nodes within the network. This attack model primarily exploits the data link layer and MAC mechanism. These components, as defined by the Open System Interconnection/International Organisation for Standardisation (OSI/ISO) model, reside in the second layer of the network architecture. The data link layer is responsible for node-to-node data transfer, a function that includes error checking and frame synchronisation. The MAC sublayer manages access to a shared medium, effectively controlling which devices are allowed to transmit data at any given time. This exploitation strategy is also pertinent to Wireless Access in Vehicular Environments (WAVE) network models and architectures, which similarly position these components at the second layer

The following characteristics inherent to this type of attack may be summarised as follows:

- **Flood**: sending REQUEST packages in short time intervals (e.g., between 5 and 10 seconds);
- **Packet storm**: sending large amounts of MAC frames to all vehicles;
- **DEAUTH**: targeting malicious MAC frames at all nodes within the scope of the vehicular network.

1) INTRUSION DETECTION ALGORITHMS

In Section III-B1, we presented the two phases of DAMASCO and the corresponding algorithms executed. Here, we provide a more detailed description of these algorithms.

Algorithm 2 is responsible for monitoring, analysing, and collecting data from the MAC frames received from the Logical Link Control (LLC) sublayer of the data link layer. This algorithm comprises a callback function, *proc_frame*, which processes each incoming frame and updates a particular data structure maintained by DAMASCO. This data structure denoted as MAC_v, stores metadata corresponding to each node in the network.

The function *proc_frame* inspects the LLC sublayer and the MAC frame to determine if it corresponds to a REQUEST (line 2). If the frame is not a REQUEST, the function returns promptly as the monitoring scope is limited to REQUEST messages exclusively. Conversely, if a REQUEST is confirmed, the function assembles a new tuple encapsulating metadata extracted from the received frame. This tuple includes a timestamp ts_{req}, the MAC and IP addresses of the sender MAC_{src} and IP_{src}, and a counter $total_{req}$ of REQUEST messages previously transmitted by the originating node. Upon completing these steps, the function can perform one of two actions:

- If MAC_v is empty (line 10) or if the MAC address of the current frame was not registered yet (line 22), it inserts the new tuple of data to the vector;
- Otherwise, the function updates the timestamp (ts_{req}, see line 17) of the node and increments a counter with the number of requests performed from that node ($total_{req}$, see line 18).

Algorithm 3 analyses the $total_{req}$ field of MAC_v (i.e., the number of REQUEST messages sent by each neighbouring vehicle) to perform anomaly detection and generate an updated reputation list. Whenever MAC_v is not empty, function *update_list* executes periodically according to a ϕ interval. This function performs the method presented in Section III-B1 to the data and fills a reputation list with the nodes corresponding outliers, classified as malicious vehicles.

As shown in Algorithm 3, the reputation list is updated periodically by analysing MAC_v. This is done locally on each vehicle within the network. The design presented so far

Algorithm 2 Monitoring, Analysis, and Collection

```
Require: $LLC_f$, $MAC_f$, $MAC_v$
 1: function $proc\_frame(LLC_f, MAC_f, MAC_v)$:
 2:     if (not $LLC_f$) or ($MAC_f$ is not $REQUEST$) then
 3:         return $MAC_v$
 4:     end if
 5:     $e \leftarrow init_{entry}()$
 6:     $e.ts_{req} \leftarrow time()$
 7:     $e.MAC_{src} \leftarrow MAC_f.MAC_{src}$
 8:     $e.IP_{src} \leftarrow MAC_f.IP_{src}$
 9:     $e.total_{req} \leftarrow 0$
10:     if $MAC_v.empty()$ then
11:         $MAC_v.append(e)$
12:     else
13:         $flag \leftarrow False$
14:         for $minMAC_v$ do
15:             if $m.MAC\_src = MAC\_src$ then
16:                 $flag \leftarrow True$
17:                 $m.ts\_req \leftarrow ts\_req$
18:                 $m.total\_req \leftarrow m.total\_req + 1$
19:                 break
20:             end if
21:         end for
22:         if not $flag$ then
23:             $MAC_v.append(e)$
24:         end if
25:     end if
26:     return $MAC_v$
27: end function
```

Algorithm 3 MAD-Based Anomaly Detection

```
Require: $MAC_v$, $ec$, $b$
function $update\_list(MAC_v, ec)$:
    $list_{rep} \leftarrow init_{list}()$
    $\mathbf{x} \leftarrow init_{vec}()$
    for $m$ in $MAC_v$ do
        $\mathbf{x}.append(m.total_{req})$
    end for
    $MAD \leftarrow b \cdot M(\|\mathbf{x} - M(\mathbf{x})\|)$
    $outliers \leftarrow find\_outliers(MAC_v, \mathbf{x}, MAD, ec)$
    for $o$ in $outliers$ do
        if ($o.total_{req} > M(\mathbf{x}) + ec \cdot MAD$) then
            $list_{rep}.append(o)$
        end if
    end for
    return $list_{rep}$
end function
```

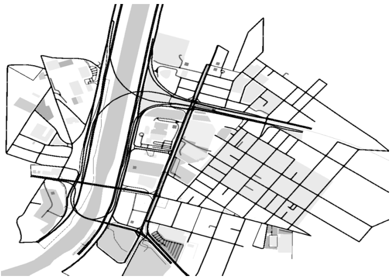

FIGURE 2. Metropolitan area of São paulo city, brazil.

implies that, after being blacklisted, a vehicle can only be removed from the reputation list if MAC_v is restarted. Other approaches may be developed to soften this protocol, such as restarting the *total_req* field of silent vehicles for a certain period. This may be implemented by monitoring the *ts_req* attribute of the vehicles in MAC_v.

IV. PERFORMANCE ANALYSIS

This section presents the procedures employed to configure the simulations used to assess the DAMASCO operation and generate a dataset for further evaluation. The following subsections thoroughly present such a methodology by outlining the specifications of the simulated environment and the dataset generation, along with the evaluation metrics employed and the results achieved.

A. SIMULATION ENVIRONMENT

We used the Network Simulator 3 (NS-3)[1] to implement an experimental environment and evaluate DAMASCO. The simulation used the IEEE 802.11p protocol stack and modelled the signal attenuation caused by different factors, such as obstacles. The attacks were represented based on a scenario generated using the Simulator for Urban MObility (SUMO) [36] and a map obtained from the OpenStreetMap tool .[2]

The use-case scenario was a fragment of São Paulo city, Brazil (Figure 2), one of the largest metropolitan areas in the world. We considered 2 km^2 of several blocks, avenues, expressways and two-way highways for vehicles moving in opposite directions. The vehicles exchanged data with each other within a radius of approximately 1 km, considering the range of the radio in the VANET and average speed between 60 to 80 km/h. In a metropolitan area, the density of vehicles is lower than in a realistic scenario where thousands of vehicles are expected per kilometre.

We use traffic vehicle densities from 120 vehicles, and the simulation starts in normal traffic conditions for a warm-up period. Within the 120 vehicles, we varied between 5%, 10%, 15%, 20%, and 25% vehicles that inject large amounts of REQUEST into the network, called malicious vehicles.

TABLE 2. Simulation parameter settings.

Parameter	Value
Simulation Time	180 s
Map area	2 km^2
Number of vehicles	120
Average velocity	$\in [60, 80]$ km^h
Malicious vehicles	$\{5\%, 10\%, 15\%, 20\%, 25\%$
Injection time	$\in [4, 5]$ s
Propagation model	Two-Ray Ground
Routing protocol	OLSR
Modulation	OFDM
Rate	6 Mbps
Datagram size (payload)	64 KB
Transport protocol	UDP
Transmission power (TX)	33.8 dBm – 60.0 dBm
Radio coverage	1 km
Number of serving vehicles	10 - 30
ϕ	2 s

The transmission process happens via broadcast, making all vehicles in the scope receive requests in short time intervals. As a result, victim vehicles typically respond to numerous requests, thus generating an unusually dense transmission. We have varied the REQUEST injection time to four and five seconds for malicious vehicles. The simulation parameter values are shown in Table 2.

B. DATASET GENERATION

The dataset was generated in this work as a product from the simulated environment. We simulated an urban environment with vehicles communicating with each other with varying proportions of malicious vehicles, as stated in the previous subsection.

By modulating different simulation parameters, we generated a dataset to capture the request and response flow. This dataset reflected both normal and abnormal traffic conditions. Normal traffic conditions were characterised by predictable volumes of request messages and responses from vehicles without any additional request packet injections. Deviations from this normal flow, specifically a surge in the number of requests, were interpreted as irregularities or abnormal flow.

Figure 3 shows the comparison between the malicious scenario (when a set of vehicles increases the sending of requests) and a standard communication scenario. For this initial analysis, we consider that 30% of vehicles are dangerous, and the values obtained are an average of REQUEST within 10-second intervals.

In a standard communication environment, we obtained an average of 231.7 requests. Considering the 30% of malicious vehicles, the simulation produced a 92% increase in the number of requests. We can observe the discrepancy of scenarios when the network is undergoing a denial of service attack, consequently allowing the classification of a typical

[1] NS-3: Network Simulator. NSNAM. https://www.nsnam.org/. Accessed: 14 of February 2023

[2] OpenStreetMap. https://www.openstreetmap.org/. Accessed: 14 of February 2023

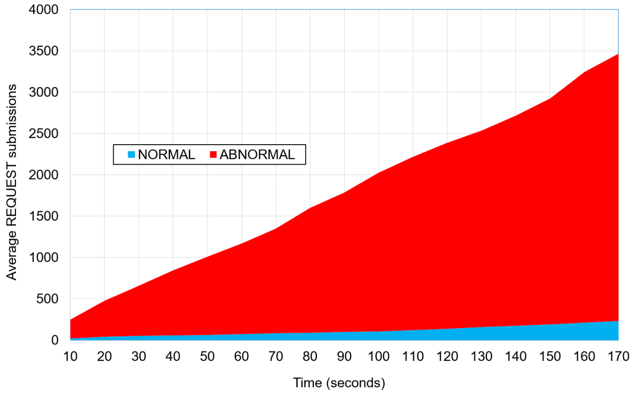

FIGURE 3. Comparison between scenarios with and without malicious vehicles.

scenario from an under-attack scenario by only analysing the outliers.

C. EVALUATION METHODOLOGY

This subsection elucidates the performance evaluation by juxtaposing the proposed DAMASCO solution with the Statistics-Based Intrusion Detection for Vanet Hosts (Sb-IDVH) [64]. We opted for Sb-IDVH as a reference due to its similarity with our work, particularly in its employment of statistical methods for attack detection and classification. Specifically, Sb-IDVH incorporates a detection and classification process that calculates the mean and adds two standard deviations.

In our experiments, we analysed various exclusion criteria ec. We considered the three exclusion criteria $ec \in \{2.0, 2.5, 3.0\}$. We named the resulting models for each ec value as DAMASCO-ec 2, DAMASCO-ec 2.5 and DAMASCO-ec 3. We also varied the certainty parameter of Sb-IDVH in Sb-IDVH-250 and Sb-IDVH 350. Furthermore, we s the DoS/DDoS attacks by increasing the number of REQUEST packages sent from a given set of nodes.

We measured the performance and efficiency of our proposed mechanism through a comparative analysis using the following base metrics utilised in this field [4], [9]:

- **True Positives (TP):** attack events correctly identified by the IDS;
- **True Negatives (TN):** normal events correctly classified by the IDS.
- **False Positives (FP):** normal events classified as attacks by the IDS;
- **False Negatives (FN):** attack events not identified by the IDS;
- **Detection rate (DR):** ratio between the number of events classified correctly and the total number of events.

The Detection Rate (DR) enables analysing the hit rate of the proposed security mechanisms considering both positive and negative ratings. Equation 3 describes the DR in which we have the sum of positive and negative detection across all classifications, including false positives and negatives.

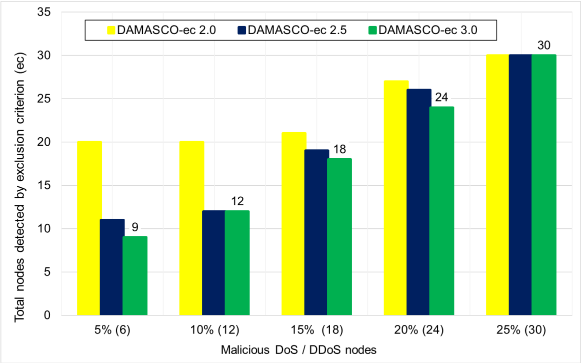

FIGURE 4. Total number of malicious nodes (DoS/DDoS) detected considering the exclusion criteria $ec \in \{2.0, 2.5, 3.0\}$.

$$DR = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

The False Positive Rate (FPR) corresponds to the ratio between the false positives and all samples classified as positive, as shown in Equation 4.

$$FPR = \frac{FP}{FP + TN} \tag{4}$$

The False-Negative Rate (FNR) corresponds to the ratio between the false negatives and all samples that were classified as negative (i.e., a request is identified as an attack, but the flow was a regular request or access), as shown in Equation 5.

$$FNR = \frac{FN}{FN + TP} \tag{5}$$

D. RESULTS

We began our analysis by comparing the overall performance of DAMASCO for the different values of exclusion criterion, considering each of the proportions of malicious nodes considered. Figure 4 presents these results. DAMASCO with $ec = 3.0$ performed best. It outperformed the DAMASCO models with $ec = 2.0$ and $ec = 2.5$ by 17% and 23%, respectively. On examining the scenario where $ec = 2.0$, it was noted that DAMASCO recorded a considerable volume of false positives, which was an outcome of its diminished precision.

As already stated, the performance metrics of DAMASCO were compared to those achieved through Sb-IDVH. Figure 5 presents the results concerning the detection rate (DR) metric. DAMASCO performed best for $ec = 2.5$ and $ec = 3.0$. With these more conservative exclusion criteria, the proposed MAD method implemented in DAMASCO yielded detection rates above 95%, compared to less than 90% for $ec = 2.0$. Regardless of the proportion of malicious nodes within the network, DAMASCO produced a higher detection rate when the exclusion criterion increased.

For the Sb-IDVH, we considered two parameter sets: the mean plus two standard deviations (SbIDVH-mean +

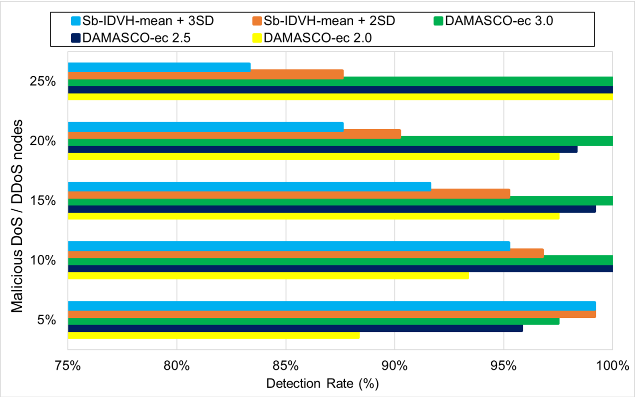

FIGURE 5. **Detection rate performance for different exclusion criteria, and considering different numbers of malicious vehicles.**

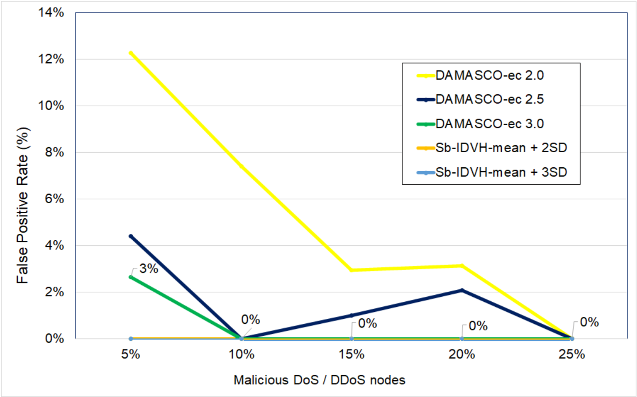

FIGURE 6. **False Positive rates for different exclusion criteria, considering different numbers of malicious vehicles.**

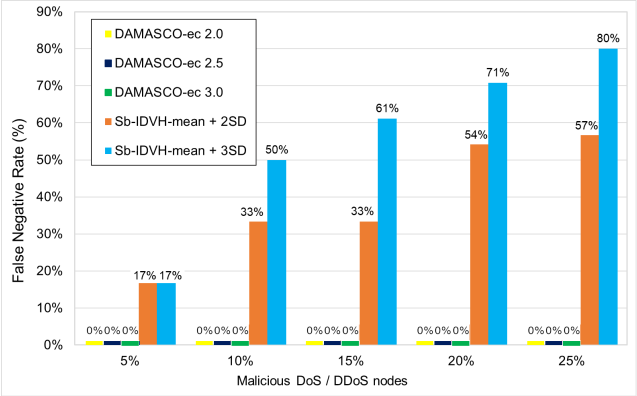

FIGURE 7. **False Negative rates for different exclusion criteria, and considering different numbers of malicious vehicles.**

2SD) and the mean plus three standard deviations (SbIDVH-mean + 3SD). When encountering 5% of malicious nodes, both configurations achieved a detection rate of 100%. However, as the proportion of malicious nodes increased, the detection rate decreased – a trend in direct contrast with the pattern observed for the DAMASCO models. Despite Sb-IDVH outperforming DAMASCO when the prevalence of malicious vehicles was as low as 5%, its detection rates fell short of those achieved by DAMASCO under all parameter configurations when this proportion increased to 10% or above.

Figure 6 highlights the false-positive rates produced. The Sb-IDVH method resulted in zero false positives across all conditions. In contrast, DAMASCO produced more false positives with less conservative exclusion criteria and lower malicious node proportions. With $ec = 2.5$ and $ec = 3.0$, the false positives were minimal.

Even with $ec \in \{2.5, 3.0\}$ and a 20% proportion of malicious vehicles, the DAMASCO threshold was not exceeded. DAMASCO threshold was not reached when considering the $ec \in \{2.5, 3.0\}$, even with a percentage of 20% of malicious vehicles. This result was achieved because the distance among the sample values was always relatively small. Therefore, the threshold of DAMASCO_ec-3 has achieved a greater distance to detect extreme values by the total number of REQUEST packages received.

Figure 7 illustrates the false negative rates for each model. DAMASCO exhibited superior performance in this metric, producing no false negatives in any of the conditions analysed. In contrast, the Sb-IDVH models resulted in a false negative rate of 17% when malicious vehicles comprised 5% of the total, and this rate escalated to as much as 57% or 80% as this proportion increased to 25%. In essence, as the proportion of malicious vehicles increases, the Sb-IDVH method's capacity to mitigate an attack diminishes.

E. DISCUSSION

Our work contributes to the field by providing a simple, yet effective approach for detecting DoS/DDoS attacks in VANETs. DAMASCO incorporates a lightweight probabilistic function to analyse and filter packets at each node within the network. This function was designed to run efficiently on nodes with limited computing power and memory, making it ideal for seamless implementation in firewalls or monitoring tools aimed at filtering and safeguarding network traffic.

We designed a proof-of-concept in which the attacks followed a normal distribution and calibrated the coefficients accordingly. When dealing with a normal distribution, the text describes that the data samples are collected from the transmission of vehicles along the path. In other words, they are the total number of requests received from neighbouring nodes that will be uniformly analysed by the MAD statistical method.

As shown in Figures 4 to 7, the performance of the anomaly detector in DAMASCO is not negatively affected when the proportion of malicious nodes in the network increase. The number of vehicles also might not affect the overall operation of the system since the whole framework operates distributedly within each node, and the algorithms for monitoring, analysis, and collection are linear with the number of requests processed by each node. Also, the integration of DAMASCO with an ITS control centre and even with vehicular cloud technologies (VANET Clouds) favours the scalability of the security mechanism, facilitating the sharing of security data, for example, the disclosure of the reputation list containing suspicious and malicious nodes.

By analysing Figure 5, we observe that the higher the proportion of malicious nodes in the network, the higher the detection rate yielded by our approach, while the Sb-IDVH model [64], based on the mean instead of the median, registered the opposite tendency (i.e., the performance was better when the proportion of malicious nodes was smaller). This phenomenon can be explained when comparing the properties of the mean/standard deviation (e.g., Sb-IDVH) to those of the median/absolute deviation (e.g., DAMASCO).

A larger proportion of outliers increases the standard deviation. Hence, a model based on mean and standard deviation presents a greater tolerance for deviant samples being considered normal. In other words, models such as the Sb-IDVH yield a higher false negative rate when the proportion of anomalous nodes rises. On the other hand, the MAD is computed based on the median of absolute deviations from the median, which makes DAMASCO robust to a large proportion of malicious nodes. However, this method can produce false positives when the proportion of malicious nodes is small.

Alternative strategies for anomaly detection that could have been considered include Bloom filters and neural networks.

Bloom filters [20], [62], [63] are probabilistic data structures that use hash functions to determine if a given input belongs to a specific set. To utilise this structure for anomaly detection, a training phase with a substantial collection of predominantly normal data would be required to calibrate the data structure. Anomalous samples could be identified by examining whether a new data point was a member of the structure.

Neural networks [12], [45], [56], particularly deep learning architectures, present a more advanced machine learning construct. They prove particularly useful when dealing with high-dimensional input data. These models are trained to learn data representations through self-supervised learning, typically using autoencoders. Samples with significant reconstruction errors are likely to be outliers, thereby facilitating anomaly detection.

Such strategies would prove advantageous if we were tackling problems that encompass intricate relationships between the input samples and the model output. This was not the case for our attack model. We modelled the dataset so that the anomaly detector operates on a single variable: the number of REQUEST packages transmitted by the vehicles. This task could be addressed with a straightforward, statistic-based model with reasonable performance, as it did in our study. Our contribution involved devising a methodology comprised of data collection from the data link layer, analysis of this data, and managing the incidents of malicious vehicles using a reputation list. This methodology is particularly fitting when implemented within V2V communication. Additionally, its simplicity facilitates deployment on devices with limited computational resources.

Concerning complexity, both DAMASCO and Sb-IDVH are linear with the number of requests. This characteristic contrasts with machine learning-based approaches, which can require maintaining a large number of parameters. Even in smaller models, the training phase would represent a bottleneck [43]. For instance, a feedforward neural network is trained through an iterative process in which each step is linear with the number of free parameters.

With the low complexity of the DAMASCO algorithm, the execution time on a desktop machine with an Intel(R) Core(TM) i7-8550U CPU @ 1.80GHz 1.99 GHz the algorithm occupied less than 0.1% of the processing capacity. If we consider that in the most current cars, such as the 2022 Tesla Model 3 and Model Y models that use AMD Ryzen (MCU3), the execution of the proposed solution on these CPUs will still be well below 0.1% of CPU usage, not impacting the execution of the proposed solution or other services embedded in the vehicle. In addition, DAMASCO uses simple structures to manipulate its information, consequently not occupying much memory space; using 16G of RAM, the system occupied less than 0.01% of memory, considering the capacity of the same models, Model 3 and Model Y 2022 Tesla model that has more than 28G of storage capacity so it would only use 0.005% of the memory.

The primary limitation of our study is its sole ability to detect DoS/DDoS attacks premised on the information flow of individual nodes since the statistical methodology hinged on the number of requests dispatched by each node within a local area. As the reputation list was confined to nodes previously encountered in the network, signalling aberrant behaviour could potentially incur delays. Moreover, a node could only be extricated from the reputation list upon system shutdown, posing complications should false positives be produced. Conversely, upon system reboot, malicious nodes would not be promptly flagged as such. Furthermore, the system does not have a mechanism to prevent trusted nodes, such as emergency vehicles, from being blacklisted.

A further limitation of our work was that our evaluation was exclusively performed through simulation. Although simulation serves as a credible method for offering a proof of concept and for assessing the system's functionality in a controlled environment, additional scrutiny employing actual data would be beneficial. Real-world environments can exhibit a multitude of behaviours that may not have been accounted for in the simulation.

The deployment of DAMASCO necessitates a V2V network with adequate radio coverage, ensuring the efficient sharing of the reputation list. According to findings from our study, implementations may employ a higher exclusion criterion, which reduces the occurrence of false positives without generating false negatives (see Figures 6 and 7). Special care must be taken if the proportion of malicious nodes is supposed to be low since this scenario is more likely to produce false positives (see Figure 7) – the most critical failure, as it can undesirably blacklist non-malicious and potentially important nodes.

V. CONCLUSION AND FUTURE WORKS

In this work, we addressed DoS/DDoS attacks performed in VANETs based on V2V communication by proposing a novel strategy in the context of ITS. While other approaches have addressed different types of attacks, most of these works rely on centralised settings or build on more sophisticated methodologies. The security mechanism proposed in this work, named DAMASCO, is based on a simple statistical model and a reputation list based on the number of REQUEST packages sent by each vehicle within a neighbourhood.

The proposed solution was evaluated through a network simulation running on a virtual environment based on a couple of blocks in São Paulo city. Different numbers of malicious vehicles were considered. We compared different parameterisations and used the Sb-IDVH method as a reference. Results showed that DAMASCO could handle the attack model, producing no false negatives, and a low false positive rate. This result differed from Sb-IDVH, which produced false negatives but was more robust to false positives.

The main advances of the solution, we can point out that DAMASCO provides a decentralised and cooperative security system for detecting DoS/DDoS attacks in VANET without any external infrastructure, such as a cellular network. In addition, the anomaly detection module is based on a straightforward statistics technique with a low impact on its execution. Therefore, the proposed method is a method with simple mathematics; any P4 programmable switch or any monitoring and packet filter software can efficiently implement it, trying to detect a possible denial attack by analysing the packet flow.

More sophisticated methods based on machine learning could be helpful if our input space had high dimensionality. Since this was not the case, and we considered only the number of REQUEST packages sent by each vehicle, we achieved high detection rates with a more straightforward technique.

In future works, we intend to enhance the simulation environment so that we can consider a more dynamic attack model. Besides, attacks other than DoS/DDoS can be considered. In this extended scenario, we can consider how the MAD method can complement other more sophisticated anomaly detection techniques, such as Bloom filters and neural networks. Moreover, to make our solution closer to a deployment case, we might integrate the DAMASCO security mechanism into a VANET Cloud and ITS control centre.

REFERENCES

[1] E. A. Shams, A. Rizaner, and A. H. Ulusoy, "Flow-based intrusion detection system in vehicular ad hoc network using context-aware feature extraction," *Veh. Commun.*, vol. 41, Jun. 2023, Art. no. 100585.

[2] C. C. Aggarwal, *Outlier Analysis*, 2nd ed. Cham, Switzerland: Springer, 2016.

[3] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *J. Netw. Comput. Appl.*, vol. 60, pp. 19–31, Jan. 2016.

[4] K. M. A. Alheeti, A. Gruebler, and K. McDonald-Maier, "Using discriminant analysis to detect intrusions in external communication for self-driving vehicles," *Digit. Commun. Netw.*, vol. 3, no. 3, pp. 180–187, Aug. 2017.

[5] G. O. Anyanwu, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "RBF-SVM kernel-based model for detecting DDoS attacks in SDN integrated vehicular network," *Ad Hoc Netw.*, vol. 140, Mar. 2023, Art. no. 103026.

[6] H. Baharlouei, A. Makanju, and N. Zincir-Heywood, "Exploring realistic VANET simulations for anomaly detection of DDoS attacks," in *Proc. IEEE 95th Veh. Technol. Conf. (VTC-Spring)*, Jun. 2022, pp. 1–7.

[7] B. Baruah and S. Dhal, "A secure road condition monitoring scheme in cloud based VANET," *Comput. Commun.*, vol. 174, pp. 131–142, Jun. 2021.

[8] A. Bicaku, M. Zsilak, P. Theiler, M. Tauber, and J. Delsing, "Security standard compliance verification in system of systems," *IEEE Syst. J.*, vol. 16, no. 2, pp. 2195–2205, Jun. 2022.

[9] J. Brownlee, *Tour of Evaluation Metrics for Imbalanced Classification*. Vermont, VIC, Australia: Imbalanced Classification, 2020.

[10] G. Buzzi-Ferraris and F. Manenti, "Outlier detection in large data sets," *Comput. Chem. Eng.*, vol. 35, no. 2, pp. 388–390, Feb. 2011.

[11] R. Chakraborty, S. Kumar, A. Awasthi, K. Suneetha, A. Rastogi, and G. Jethava, "Machine learning based novel frameworks developments and architectures for secured communication in VANETs for smart transportation," *Soft Comput.*, vol. 27, pp. 1–11, May 2023.

[12] R. Chalapathy, N. L. D. Khoa, and S. Chawla, "Robust deep learning methods for anomaly detection," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 3507–3508.

[13] A. L. Cristiani, D. D. Lieira, R. I. Meneguette, and H. A. Camargo, "A fuzzy intrusion detection system for identifying cyber-attacks on IoT networks," in *Proc. IEEE Latin-Amer. Conf. Commun. (LATINCOM)*, Nov. 2020, pp. 1–6.

[14] N. D. V. Dalarmelina, M. A. Teixeira, and R. I. Meneguette, "A real-time automatic plate recognition system based on optical character recognition and wireless sensor networks for ITS," *Sensors*, vol. 20, no. 1, p. 55, Dec. 2019.

[15] N.-N. Dao, T. V. Phan, U. Sa'ad, J. Kim, T. Bauschert, D.-T. Do, and S. Cho, "Securing heterogeneous IoT with intelligent DDoS attack behavior learning," *IEEE Syst. J.*, vol. 16, no. 2, pp. 1974–1983, Jun. 2022.

[16] G. de C. Bertoli, L. A. P. Junior, O. Saotome, and A. L. dos Santos, "Generalizing intrusion detection for heterogeneous networks: A stacked-unsupervised federated learning approach," *Comput. Secur.*, vol. 127, Apr. 2023, Art. no. 103106.

[17] L. H. de Melo, G. de C. Bertoli, L. A. Pereira, O. Saotome, M. F. Domingues, and A. L. dos Santos, "Generalizing flow classification for distributed denial-of-service over different networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022, pp. 879–884.

[18] F. Dutra, K. Bonfim, C. Siqueira, L. A. Pereira, A. Santos, and R. I. Meneguette, "DISMISS-BSM: An architecture for detecting position spoofing in basic safety messages," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2022, pp. 3381–3386.

[19] G. P. R. Filho, R. I. Meneguette, G. Maia, G. Pessin, V. P. Gonçalves, L. Weigang, J. Ueyama, and L. A. Villas, "A fog-enabled smart home solution for decision-making using smart objects," *Future Gener. Comput. Syst.*, vol. 103, pp. 18–27, Feb. 2020.

[20] S. Garg, A. Singh, G. S. Aujla, S. Kaur, S. Batra, and N. Kumar, "A probabilistic data structures-based anomaly detection scheme for software-defined Internet of Vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3557–3566, Jun. 2021.

[21] M. R. Habibi, H. R. Baghaee, F. Blaabjerg, and T. Dragicevic, "Secure control of DC microgrids for instant detection and mitigation of cyber-attacks based on artificial intelligence," *IEEE Syst. J.*, vol. 16, no. 2, pp. 2580–2591, Jun. 2022.

[22] S. Hamdan, A. Hudaib, and A. Awajan, "Detecting Sybil attacks in vehicular ad hoc networks," *Int. J. Parallel, Emergent Distrib. Syst.*, vol. 36, no. 2, pp. 69–79, Mar. 2021.

[23] F. R. Hampel, "The influence curve and its role in robust estimation," *J. Amer. Stat. Assoc.*, vol. 69, no. 346, pp. 383–393, Jun. 1974.

[24] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, "VANet security challenges and solutions: A survey," *Veh. Commun.*, vol. 7, pp. 7–20, Jan. 2017.

[25] A. Haydari and Y. Yilmaz, "RSU-based online intrusion detection and mitigation for VANET," *Sensors*, vol. 22, no. 19, p. 7612, Oct. 2022.

[26] P. J. Huber, ''Robust statistics,'' in *International Encyclopedia of Statistical Science*. Berlin, Germany: Springer, 2011, pp. 1248–1251.

[27] N. H. Hussein, C. T. Yaw, S. P. Koh, S. K. Tiong, and K. H. Chong, ''A comprehensive survey on vehicular networking: Communications, applications, challenges, and upcoming research directions,'' *IEEE Access*, vol. 10, pp. 86127–86180, 2022.

[28] L. J. Vinita and V. Vetriselvi, ''Federated learning-based misbehaviour detection on an emergency message dissemination scenario for the 6G-enabled Internet of Vehicles,'' *Ad Hoc Netw.*, vol. 144, May 2023, Art. no. 103153.

[29] S. A. Jan, N. U. Amin, M. Othman, M. Ali, A. I. Umar, and A. Basir, ''A survey on privacy-preserving authentication schemes in VANETs: Attacks, challenges and open issues,'' *IEEE Access*, vol. 9, pp. 153701–153726, 2021.

[30] R. N. Kamoi, L. A. P. Júnior, F. A. N. Verri, C. A. C. Marcondes, C. H. G. Ferreira, R. I. Meneguette, and A. M. D. Cunha, ''Platoon grouping network offloading mechanism for VANETs,'' *IEEE Access*, vol. 9, pp. 53936–53951, 2021.

[31] K. Kim, J. S. Kim, S. Jeong, J.-H. Park, and H. K. Kim, ''Cybersecurity for autonomous vehicles: Review of attacks and defense,'' *Comput. Secur.*, vol. 103, Apr. 2021, Art. no. 102150.

[32] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, ''DDoS in the IoT: Mirai and other botnets,'' *Computer*, vol. 50, no. 7, pp. 80–84, 2017.

[33] S. Kumar and K. Dutta, ''Intrusion detection in mobile ad hoc networks: Techniques, systems, and future challenges,'' *Secur. Commun. Netw.*, vol. 9, no. 14, pp. 2484–2556, Sep. 2016.

[34] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, ''Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median,'' *J. Exp. Social Psychol.*, vol. 49, no. 4, pp. 764–766, Jul. 2013.

[35] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, ''Federated learning: Challenges, methods, and future directions,'' *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50–60, May 2020.

[36] P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wiessner, ''Microscopic traffic simulation using SUMO,'' in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2575–2582.

[37] G. Loukas, Y. Yoon, G. Sakellari, T. Vuong, and R. Heartfield, ''Computation offloading of a vehicle's continuous intrusion detection workload for energy efficiency and performance,'' *Simul. Model. Pract. Theory*, vol. 73, pp. 83–94, Apr. 2017.

[38] N. T. Luong, A. Q. Nguyen, and D. Hoang, ''FAPDRP: A flooding attacks prevention and detection routing protocol in vehicular ad hoc network using behavior history and nonlinear median filter transformation,'' *Wireless Netw.*, pp. 1–28, Oct. 2022.

[39] N. Lyamin, A. Vinel, M. Jonsson, and J. Loo, ''Real-time detection of denial-of-service attacks in IEEE 802.11p vehicular networks,'' *IEEE Commun. Lett.*, vol. 18, no. 1, pp. 110–113, Jan. 2014.

[40] R. Meneguette, R. De Grande, J. Ueyama, G. P. R. Filho, and E. Madeira, ''Vehicular edge computing: Architecture, resource management, security, and challenges,'' *ACM Comput. Surv.*, vol. 55, no. 1, pp. 1–46, Jan. 2023.

[41] R. I. Meneguette and A. Boukerche, ''A cooperative and adaptive resource scheduling for vehicular cloud,'' in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jul. 2017, pp. 398–403.

[42] R. I. Meneguette and A. Boukerche, ''Vehicular clouds leveraging mobile urban computing through resource discovery,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 6, pp. 2640–2647, Jun. 2020.

[43] T. M. Mitchell, *Machine Learning*, vol. 1. New York, NY, USA: McGraw-Hill, 1997.

[44] G. Muruti, F. A. Rahim, and Z. bin Ibrahim, ''A survey on anomalies detection techniques and measurement methods,'' in *Proc. IEEE Conf. Appl., Inf. Netw. Secur. (AINS)*, Nov. 2018, pp. 81–86.

[45] S. Naseer, Y. Saleem, S. Khalid, M. K. Bashir, J. Han, M. M. Iqbal, and K. Han, ''Enhanced network anomaly detection based on deep neural networks,'' *IEEE Access*, vol. 6, pp. 48231–48246, 2018.

[46] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, ''Machine learning for anomaly detection: A systematic review,'' *IEEE Access*, vol. 9, pp. 78658–78700, 2021.

[47] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, ''Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on Internet-scale IoT exploitations,'' *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2702–2733, 3rd Quart., 2019.

[48] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, ''Deep learning for anomaly detection: A review,'' *ACM Comput. Surv. (CSUR)*, vol. 54, no. 2, pp. 1–38, 2021.

[49] A. Paranjothi and M. Atiquzzaman, ''A statistical approach for enhancing security in VANETs with efficient rogue node detection using fog computing,'' *Digit. Commun. Netw.*, vol. 8, no. 5, pp. 814–824, Oct. 2022.

[50] R. Pereira, A. Boukerche, M. A. C. da Silva, L. H. V. Nakamura, H. Freitas, G. P. R. Filho, and R. I. Meneguette, ''FORESAM—FOG paradigm-based resource allocation mechanism for vehicular clouds,'' *Sensors*, vol. 21, no. 15, p. 5028, Jul. 2021.

[51] R. S. Pereira, D. D. Lieira, M. A. C. da Silva, A. H. M. Pimenta, J. B. D. da Costa, D. Rosário, and R. I. Meneguette, ''A novel fog-based resource allocation policy for vehicular clouds in the highway environment,'' in *Proc. IEEE Latin-Amer. Conf. Commun. (LATINCOM)*, Nov. 2019, pp. 1–6.

[52] P. J. Rousseeuw and C. Croux, ''Alternatives to the median absolute deviation,'' *J. Amer. Stat. Assoc.*, vol. 88, no. 424, pp. 1273–1283, Dec. 1993.

[53] F. Sakiz and S. Sen, ''A survey of attacks and detection mechanisms on intelligent transportation systems: VANETs and IoV,'' *Ad Hoc Netw.*, vol. 61, pp. 33–50, Jun. 2017.

[54] C. Sanders and J. Smith, *Applied Network Security Monitoring: Collection, Detection, and Analysis*. Amsterdam, The Netherlands: Elsevier, 2013.

[55] D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, and D. Dennison, ''Hidden technical debt in machine learning systems,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 1–9.

[56] B. Sharma, L. Sharma, and C. Lal, ''Anomaly detection techniques using deep learning in IoT: A survey,'' in *Proc. Int. Conf. Comput. Intell. Knowl. Economy (ICCIKE)*, Dec. 2019, pp. 146–149.

[57] S. Sharma and A. Kaul, ''A survey on intrusion detection systems and honeypot based proactive security mechanisms in VANETs and VANET cloud,'' *Veh. Commun.*, vol. 12, pp. 138–164, Apr. 2018.

[58] B. Shivanand, S. S. Tangade, G. D. Devanagavi, and S. S. Manvi, ''A survey on security and safety in vehicular ad hoc networks (VANETs) cloud,'' in *Cognitive Informatics and Soft Computing*. Singapore: Springer, 2021, pp. 309–319.

[59] S. So, J. Petit, and D. Starobinski, ''Physical layer plausibility checks for misbehavior detection in V2X networks,'' in *Proc. 12th Conf. Secur. Privacy Wireless Mobile Netw.*, May 2019, pp. 84–93.

[60] W. Stallings and T. Case, *Redes e Sistemas de Comunicação de Dados*. Rio de Janeiro, Brazil: Elsevier, 2016.

[61] X. Tang, C. Guo, Y. Ren, C. Wang, and K. R. Choo, ''A global secure ranked multikeyword search based on the multiowner model for cloud-based systems,'' *IEEE Syst. J.*, vol. 16, no. 2, pp. 1717–1728, Jun. 2022.

[62] P. Xiao, Z. Li, H. Qi, W. Qu, and H. Yu, ''An efficient DDoS detection with Bloom filter in SDN,'' in *Proc. IEEE Trustcom/BigDataSE/ISPA*, Aug. 2016, pp. 1–6.

[63] R. Xiao, Z. Zeng, C. Xiao, and S. Zhang, ''IIoT hidden anomaly detection based on locality sensitive Bloom filter,'' *J. Comput. Appl.*, vol. 41, no. 12, pp. 3620–3625, 2021.

[64] K. Zaidi, M. B. Milojevic, V. Rakocevic, A. Nallanathan, and M. Rajarajan, ''Host-based intrusion detection for VANETs: A statistical approach to rogue node detection,'' *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6703–6714, Aug. 2016.

EDIVALDO PASTORI VALENTINI received the master's degree in computer science from the Institute of Biosciences, Letters and Exact Sciences, State University Estadual Paulista ''Júlio de Mesquita Filho,'' in 2020. He is currently a Specialist in internet systems design with the Faculty of Philosophy, Sciences and Letters of Catanduva, in 2005, and in Pedagogical Training for High School Professional Education, State Center for Technological Education Paula Souza de São Paulo, in 2018. He is currently a Full Professor in computing with the Federal Institute of Education, Science and Technology of São Paulo (IFSP). He lives learning and researching security, vehicle networks, intelligent transport systems, intrusion detection systems, the Internet of Things, and education for the safe use of the internet.

GERALDO PEREIRA ROCHA FILHO received the master's and Ph.D. degrees in computer science and computational mathematics from ICMC-USP with a FAPESP Scholarship. He is currently a Professor with the Department of Exact and Technological Sciences, State University of Southwest Bahia (UESB). He was an effective Professor with the Computer Science Department, University of Brasília (UnB) (2019–2022). In 2022, he requested a vacancy with UnB to UESB. He was a Researcher with the Institute of Computing, UNICAMP, through the postdoctorate funded by FAPESP. In the last five years, he has obtained more than 24 publications in international journals and more than 32 publications in conferences. His research interests include wireless sensor networks, vehicular networks, smart grids, smart home, and machine learning.

ROBSON EDUARDO DE GRANDE (Member, IEEE) received the Ph.D. degree in computer science from the University of Ottawa, Canada, in 2012. He is currently an Associate Professor with the Department of Computer Science, Brock University, Canada. His research interests include large-scale distributed and mobile systems, cloud computing, performance modeling and simulation, computer networks, vehicular networks, intelligent transportation systems, and distributed simulation systems, actively contributing to these areas. He has served as a Technical Program and the Special Session Co-Chair for several IEEE and ACM-sponsored conferences, including IEEE/ACM DS-RT, ACM MobiWac, ACM DIVANet, and IEEE DCOSS International Workshop on Urban Computing.

CAETANO MAZZONI RANIERI received the Graduate degree in computer science from São Paulo State University (UNESP), in 2013, and the master's and Ph.D. degrees from the Institute of Mathematical and Computer Sciences, University of São Paulo (ICMC-USP), in 2016 and 2021, respectively. During the Ph.D. degree, he was a Visiting Scholar with Heriot-Watt University, Scotland, in 2020. Currently, he is a Postdoctoral Research Fellow of ICMC-USP, with research focused on artificial intelligence in the context of the Internet of Things. Has experience in machine learning, deep learning, neurorobotics, and human–robot interaction. His research interests include classifying human behavior based on multimodal signals and its replication on biologically-inspired robotics systems.

LOURENÇO ALVES PEREIRA JÚNIOR received the B.Sc. degree (Hons.) in computer science from the University of Alfenas (UNIFENAS), in 2006, and the M.Sc. and Ph.D. degrees in computer science and computational mathematics from the University of Sao Paulo (ICMC/USP), in 2010 and 2016, respectively. He is currently an Assistant Professor with the Department of Computer Systems, Brazilian Aeronautics Institute of Technology—ITA. He is also the Head of the Laboratory of Command & Control and a Cyber Defense Member. His research interests include computer networks and cybersecurity. He has also acted as a Reviewer of *IEEE Communications Magazine*, IEEE ACCESS, IEEE LATIN AMERICA TRANSACTIONS, *JNSM*, *Vehicular Communications* (Elsevier), *ACM TOIT*, IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, and IEEE TRANSACTIONS ON MACHINE LEARNING IN COMMUNICATIONS AND NETWORKING.

RODOLFO IPOLITO MENEGUETTE (Member, IEEE) received the bachelor's degree in computer science from Paulista University (UNIP), Brazil, in 2006, the master's degree from the Federal University of São Carlos (UFSCar), in 2009, and the Ph.D. degree from the University of Campinas (Unicamp), Brazil, in 2013. He is currently a Professor with the University of São Paulo (USP). In 2017, he did his postdoctorate with the PARADISE Research Laboratory, University of Ottawa, Canada. His research interests include vehicular networks, resources management, flow of mobility, and vehicular clouds.