Düzce University
Journal of Science & Technology

Research Article

A Comprehensive Survey On Machine Learning-Based Intrusion Detection System for Vehicular Area Network Architectures

Deniz BALTA[a]*, Ünal ÇAVUŞOĞLU [a], Musa BALTA[b]

[a] *Yazılım Mühendisliği, Bilgisayar ve Bilişim Bilimleri Fakültesi, Sakarya Üniversitesi, Sakarya, TÜRKİYE*
[b] *BilgisayarMühendisliği, Bilgisayar ve Bilişim Bilimleri Fakültesi, Sakarya Üniversitesi, Sakarya, TÜRKİYE*
** Sorumlu yazarın e-posta adresi: ddural@sakarya.edu.tr*

DOI: 10.29130/dubited.1372131

ABSTRACT

Today, the development of communication technologies causes changes in many different areas. One of these areas is VANET (Vehicular Area Network) application area. With the increase in usage areas in the VANET field, ensuring VANET network security has become more critical. Many different systems have been developed to detect attacks on VANET networks. Machine learning-based systems are one of the most widely used methods in developing these intrusion detection systems (IDS). In this article, research on machine learning-based VANET IDS, which has been done recently in the literature, has been carried out. First, VANET architecture and security requirements are presented, then a comprehensive literature summary is given, and comparisons are made on different parameters. As a result, it has been determined that many different machine learning models are used in IDSs and perform high-performance detection. In addition to the machine learning algorithm used in the performance of IDSs, it has been shown that many different parameters play a critical role in the performance. The paper aims to guide new studies in this field with the gains that will increase the performance of intrusion detection systems because of the literature comparison (considering criteria such as machine learning model, simulation tools, dataset, machine learning algorithm, and performance criteria).

Keywords: *IDS, Machine learning, Security, Vehicular networks*

Araçsal Ağ Mimarisi İçin Makine Öğrenmesi Tabanlı Saldırı Tespit Sistemi Üzerine Kapsamlı Bir Araştırma

Öz

Günümüzde iletişim teknolojilerinin gelişmesi birçok farklı alanda değişimlere neden olmaktadır. Bu alanlardan biri de VANET (Araç Alan Ağı) uygulama alanıdır. VANET alanında kullanım alanlarının artmasıyla birlikte VANET ağ güvenliğinin sağlanması daha kritik hale gelmiştir. VANET ağlarına yapılan saldırıları tespit etmek için birçok farklı sistem geliştirilmiştir. Makine öğrenimi tabanlı sistemler, bu saldırı tespit sistemlerinin (STS-Intrusion Detection Systems IDS) geliştirilmesinde en yaygın kullanılan yöntemlerden biridir. Bu makalede literatürde son dönemde yapılan makine öğrenmesi tabanlı VANET IDS üzerine araştırmalar yapılmıştır. Öncelikle VANET mimarisi ve güvenlik gereksinimleri sunulmuş, ardından kapsamlı bir literatür özeti verilmiş ve farklı parametreler üzerinden karşılaştırmalar yapılmıştır. Sonuç olarak, saldırı tespit sistemlerinde birçok farklı makine öğrenmesi modelinin kullanıldığı ve yüksek performanslı tespit gerçekleştirdiği tespit edilmiştir. STS'nin performansında kullanılan makine öğrenmesi algoritmasının yanı sıra birçok farklı parametrenin de performansta kritik rol oynadığı gösterilmiştir. Makale, literatür karşılaştırması (makine öğrenme modeli, simülasyon araçları,

Received: 06/10/2023, Revised: 15/10/2023, Accepted: 07/11/2023

veri seti, makine öğrenme algoritması ve performans kriterleri gibi kriterler dikkate alınarak) sayesinde saldırı tespit sistemlerinin performansını artıracak kazanımlarla bu alanda yapılacak yeni çalışmalara rehberlik etmeyi amaçlamaktadır.

Anahtar Kelimeler: *Saldırı tespit sistemleri, Makine öğrenmesi, Güvenlik, Araçsal ağlar*

I. INTRODUCTION

According to the statistics in The Global status report about road safety published by the World Health Organization (WHO) in 2018, it is emphasized that the number of people who lost their lives in traffic accidents worldwide reached 1.35 million every year [1]. Although serious measures have been taken to reduce this death rate in many countries by imposing strict traffic rules or establishing road-wide monitoring systems, the expected decrease has not been achieved. Due to the serious increase in important issues such as traffic accidents, fuel consumption, road congestion, and environmental pollution, many researchers have suggested studies on Vehicular ad hoc network (VANET) architectures to find solutions to these problems in recent years [2-7]. In this network paradigm, in which vehicles are modeled as mobile nodes by communicating with each other or roadside units, studies are carried out in the field of traffic efficiency, traffic safety, and infotainment.

VANET systems can include vehicles, drivers, pedestrians, and peripherals such as remote servers and roadside units (RSU). Each item has many hardware or software units such as OBUs, computing units, microprocessors, wireless adapters, data storage HMI, sensors, actuators, bus systems and other network components. According to architectural models, the integration is determined, and the communication is provided with data exchange between VANET nodes [8, 9, 10]. If a critical situation occurs, it is important to deliver a high level of communication with a low error rate. As noted above, VANETs can have numerous hardware and software-integrated snap-ins. The integrated structures of these components may cause security vulnerabilities in communication infrastructures between vehicles, between vehicles and infrastructure, or between infrastructure and cloud systems. These security vulnerabilities encountered on VANET architectures in academic studies generally include scalability, limited resources, dynamic topology changes, transmission quality, device discovery, and bandwidth optimization [8-14]. In the last few years, it has become one of the most studied subjects of intelligent transportation systems, as security deficiencies can cause fatal consequences in VANET systems, especially regarding traffic safety.

In VANET architectures, a malicious message intervening during the communication between the nodes may interrupt the whole system and cause the services to stop. As a result of a cyber-attack, a roadside unit may not be able to serve the vehicles in the field and the center [11, 12]. In addition, the content of the messages in the control channel can be changed. Considering very different security scenarios such as these, the importance of intrusion detection systems is better understood to protect VANET systems against malicious activities and detect attacks in advance. Intrusion detection systems to be developed for these systems should be modeled using artificial intelligence algorithms, unlike rule-defined structures, to be effective against known or unknown cyber-attacks such as man-in-the-middle, denial of service, fraud, and information manipulation [10-14]. In this context, in this study, the types of attacks on different layers of VANET architectures (physical, transfer, application, etc.) and application areas (traffic security, traffic efficiency, infotainment) and the existing security solutions in the literature to prevent these attacks are examined in detail. Scientific contributions of the study are listed below:

- As a result of the examined academic studies, it has been seen that there are many survey studies on security issues on VANET architectures. However, in these studies, it has been determined that the artificial intelligence-based IDS models proposed for the security of VANET systems have not been analyzed in depth with the layered structures and application areas of today’s VANET architectures. For this reason, in this study, the machine learning-based VANET IDS systems proposed in the literature, especially under variable traffic conditions and different cyber threats, are extensively compared by considering the concepts of quality of service such as flexibility, scalability, and accuracy.

- In the study, machine learning-based VANET IDS systems that have been made recently in the literature were examined, and comparisons were made using many different parameters (machine learning model, simulation tools, dataset, machine learning algorithm, and performance criteria). As a result of these comparisons, inferences were made to increase the intrusion detection systems' performance. It is thought that the results obtained will lead to new studies in this field.

In the following parts of the study, after the introduction, in Chapter 2, basic information about the architectural structures and security requirements of VANET systems will be given. In Chapter 3, the methods, datasets and performance criteria used by machine learning-based VANET intrusion detection systems in the recent literature will be extensively analyzed and these studies will be compared. In Chapter 4, an evaluation of the results obtained will be presented.

II. VANET ARCHITECTURE AND SECURITY REQUIREMENTS

Recently, with the developments in the vehicle industry and wireless communication technologies, vehicle network systems, which are a subset of ad hoc mobile networks, have begun to be developed, in which vehicles communicate with each other, especially to implement traffic applications related to vehicles, passengers, pedestrians and drivers. To ensure the quality of service in the communication between vehicles and roadside units in urban traffic, vehicular networks that can adapt to rapid network topology changes, support multiple or single hop transmission, provide high data rates, and provide solutions to different traffic applications are preferred instead of MANET architecture. Vehicular networks can be examined under 3 main application titles within the scope of the objectives and targets determined by the international standards and directives of intelligent transportation systems [10, 5, 6]:

- ***Traffic safety applications:*** It deals with the applications that aim to reduce the rate of traffic accidents and accordingly, reduce the probability of injury and death. These are the most studied within the scope of inter-vehicle communication.

- ***Traffic efficiency applications:*** This class of applications focuses on improving the traffic coordination, traffic flow, and traffic assistance of vehicles based on local information, messages, and maps defined by time and place.

- ***Infotainment applications:*** These applications offer convenience and comfort to the driver and passengers. For this purpose, it has an application scope to support all message types that offer entertainment and useful messages. In general, the classification tree in Figure 1 has been prepared for relating these application areas in vehicular networks with smart transportation system applications in the literature and in practice.

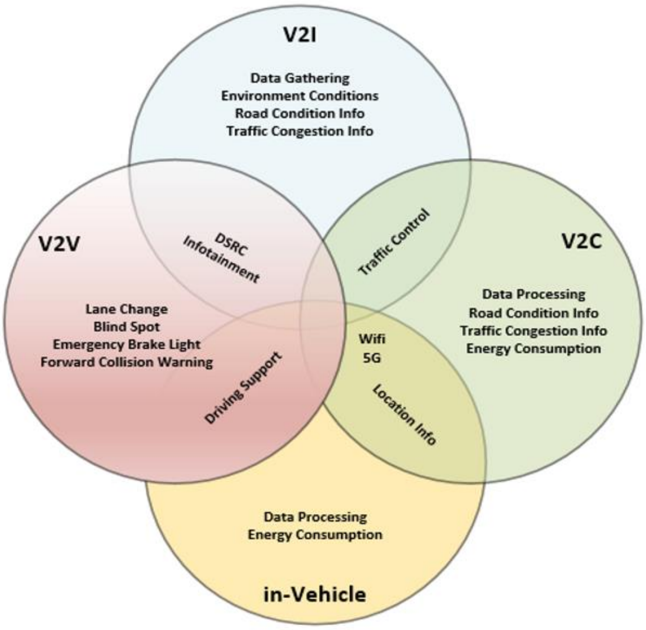

Figure 1. *VANET architecture application areas*

Vehicular networks support two different communication systems, in-vehicle and out-of-vehicle. Today's vehicle structures have become complex distributed computer systems that provide various requests and functions through communication technologies. This complex structure has subsystems with different control and communication characteristics such as chassis, power transmission, safety, body, and comfort electronics. To meet the communication (in-vehicle) requirements of these subsystems, multiple bus systems such as CAN, LIN, FlexRay and MOST can be used in the vehicle. For non-vehicle communications in vehicular networks, communication-based on DSRC standards and 802.11p protocol is used, consisting of 7 channels, each of which provides a data rate of 6-27 Mbps, with a 75 MHz bandwidth in the 5.9 GHz band. 6 of these 7 channels are service channels. They are used for traffic efficiency and infotainment applications, while 1 channel is allocated as a control channel for traffic safety applications [2-6].

A. VEHICULAR NETWORK ARCHITECTURES

While technological developments in sub-systems continue, communication between vehicles, roadside units, or cloud systems should also be provided within the scope of the application areas of instrumental networks mentioned above. WAVE, CALM and C2CNet peer-to-peer vehicle network architectures have been developed to implement these systems. Although these architectures have developed different protocol proposals, especially in the 3rd, 4th, 5th, 6th and 7th layers, according to the layered architecture to implement safety, infotainment applications, and traffic efficiency, they have agreed on the 802.11p protocol in the 5.9 GHz band in the 1st and 2nd layer. WAVE (Wireless Access in Vehicular Environment) architecture was initiated by the USA in 2004 as a project of intelligent transportation systems based on vehicular networks. The WAVE architecture, which consists of IEEE 1609 protocol cluster, consists of 802.11p, a different version of IEEE 802.11a, with OFDM (Orthogonal Frequency Division Multiplexing) mechanism that can support different data rates, determined according to the coding rate and modulation type in its physical layer. CALM architecture is an ISO-recommended instrumental network architecture, derived from the COOPERS and SAFESPOT projects, which are European consortium smart transportation system implementations. The CALM architecture, which has a 3-layer structure as Application, Network, and Interface layer, uses DSRC standards in the 5.9 GHz band as in the WAVE architecture. The CALM architecture, which supports infrared communication for short-distance communications, also supports GSM and UTMS technologies in the interface layer for long-distance communications [2-4]. This architecture consists of 3 main components:

1. CALM interface manager; It displays and records the status of each communication interface, which helps in decision-making along with the channel quality.

2. CALM network manager; It manages transfers to alternative media.

3. CALM application manager; it manages application transmission requirements. It interacts with the communication interfaces to obtain information about available environments, sending commands to the network administrator to establish a connection.

The car-to-car consortium aims to create an open-source European industry standard and develop active traffic safety practices. In this context, the C2CNet architecture, created by the European vehicle industry, defined the C2CNet protocol as different from the IP protocol. This protocol is designed to support both security and non-security applications. This architectural structure, which uses location-based algorithms for routing, uses a 30 MHz bandwidth for security applications in the 5.9 GHz band. The network layer supports multi-hop communication based on geo-addressing and routing. We can list the general features of the C2CNet architecture.

1. Fast data communication in vehicle-to-vehicle and vehicle-to-infrastructure communications

2. Support for forwarding messages that include infotainment and security messages.

3. Support for IEEE 802.11p for short-range wireless LAN technologies, traditional wireless LAN technologies (IEEE 802.11a/b/g/n) and radio technologies for long-distance communication (UMTS and GPRS) Comparisons of these architectures proposed for vehicular networks based on application, communication, routing, and technology are given in Table 1.

Table 1. *Comparative table of VANET architectures*

Parameter/ Protocol Criteria	**C2CNet**	**CALM**	**WAVE**
Vendor	Car-car consortium	ISO	ABD
Feature	Multi-hop and geo-routing	Multi-transmission media support (802.11p, DSRC, W-LAN)	802.11p for emergency mess. on Mac layer
Scope	Traffic Safety	Traffic efficiency and Infotainment	All
Addressing method	Geo-routing	IP addressing	IP addressing
Routing	Mac Protocol+IPv6	Mobile IPv6	Different channels +IPv6
Hopping method	Single&Multi hopping	Single hopping	Single hopping
Communication mode	Unicast, broadcast, geo-unicast, geo-broadcast	Unicast, broadcast	Unicast
Simulator	Commercial	Commercial	Open source

B. SECURITY REQUIREMENTS AND ATTACK TYPES IN VANETS

B. 1. Security Requirements

VANET systems have certain security requirements for secure communication like other networks. These systems' security requirements are defined as authentication, confidentiality, availability, non-repudiation and integrity [8-14, 58].

Authentication: All messages from the source to the destination node must be authenticated during VANET communication. There are many ways to authenticate the message in VANET [57, 59]. Key management is one of these methods. With this method, tools will assign a private key to each message to ensure secure communication. After the message is received by the target vehicle, the accuracy of the key is checked. The most well-known way to provide security in this way is to use a digital signature. In this method, signing each message with Elliptic Curve Encryption (ECC) provides an efficient solution for vehicles and is especially preferred against Sybil attacks.

Confidentiality: During data transfer between RSU and RSU or between RSU and OBU, messages may include private information and contact information [9, 10]. This information may be the personal information of drivers such as his license, name, age or data such as the vehicle's travel path (route), speed. Temporary keys placed in the TPD (Tamper-Proof Device) can be used to resolve this issue by periodically changing them. Additionally, ELP (Electronic License Plate) can be used to hide the driver's identity.

Availability; The real-time VANET systems inherently vulnerable to Sybil and DoS attacks. Fast messaging during communication is not secure enough. So an attacker can shrink all the data. This availability issue in VANET can be resolved by increasing the message size in the source node or fragmentation of messages in the middle nodes. Also another solution for availability is using an efficient routing protocol (like AODV-Ad hoc on Demand Vector) [8, 9].

Integrity: With this method, it is checked whether the messages have changed during communication. Because altered or incorrect data can cause traffic jams or car accidents. For example, the coordinates of an accident on the highway can be changed by an attacker, whereupon other cars can change their speed and direction due to this. This can endanger drivers.

Non-repudiation: It will provide a control mechanism to reveal the attacker after an attack. With these applications, it is aimed at maintaining communication, collect evidence and providing usable space against VANET crimes. Though non-rejection, TPD saves id, speed, direction, etc. inside the vehicle. Real-time; Real-time constraints in-vehicle systems are the most important thing. Because of the high mobility of the node, communication between nodes can be easily interrupted. For real-time target-to-source response, unfortunately, real-time sometimes put security on the back burner.

B.2. Well-Known Attack Types for VANET Networks

VANET systems, like other networks, can face a variety of attacks. However, there are some differences of VANET according to other networks because of the types of attacks that are vary based on VANET's features such as limited bandwidth, dynamic topology changes and real-time restrictions. VANET is a subset of MANET, but their attack work is not the same but similar. The most well-known attacks on VANET systems are described below and in Figure 2 [12, 13, 14].

Denial of Service Attacks: This attack is the most common type of malicious attack. DoS attacks have two main purposes. The first purpose is to consume the bandwidth of the communication environment, and the second is to prevent the vehicles from accessing the network services. Two crucial and dangerous DoS attack types are Wormhole and Blackhole attacks in VANET.

Wormhole attacks: After a high-speed connection is established between two remote nodes, legitimate tools in the transmission area of the nodes (X and Y) use this connection to transmit their data. An attacker can drop data over the connection.

Blackhole attacks: In the communication medium, a malicious node presents itself as a central node and then drops packets.

Sybil Attacks: This type of attack aims to confuse other vehicles. Many aliases are created by the attackers. Due to these fake messages, vehicles must change their direction or speed and unwanted traffic situations may arise.

Eavesdropping Attacks: It is a type of attack used against privacy. For these attacks to be successful, the attacker must be near the RSU or in a vehicle, eavesdropping on the crash data and communications medium. To prevent this attack, message encryption can be preferred.

Impersonation Attacks: In these attacks, the attackers keep their vehicle identities secret and pretend to be another vehicle. Using IP and MAC spoofing, the attacker obtains the identity of the permissive users. When they get the ID, the attackers can send the wrong message such as changing the coordinates of the traffic accident. Such attacks can be prevented by using the certificate system.

Alteration Attacks: The communication between the vehicles or the vehicle RSU is listened to by the attackers, and they can change the data as he wishes when he finds the information available to them.

Replay Attacks: The attacker gets the initial transmission packets when connecting two vehicles. VANET architecture does not prevent such attacks. These attacks aim to consume bandwidth.

Location Falsification: The attacker modifies the coordinates of the traffic accident on the highway or in the city center. When an attacker sees a car accident, fake GPS (Global Positioning System) coordinates are broadcasted. And then, other vehicles change direction to the new fake coordinate and suddenly a traffic accident or jam can occur. These attacks against VANET systems can be performed by different attackers. Attacker profiles are defined as:

From outside and inside, Outside attackers are unauthenticated nodes in the vehicle system. Usually, they are located near the RSU. Then they constantly listen to the communication medium to obtain information. The inside attacker is the authenticated nodes in the tool system, acting like a free tool until the attack. Active and passive; Broadcast messages are sent by active attackers to harm other nodes. But passive attackers don’t send messages regularly, on the contrary, they wait for the right time to attack. Malicious and rational; The rational attacker is careful during attacks and defines a specific victim. But malicious attackers do not have a specific target, they can randomly attack the communication medium or any node using DoS attacks.

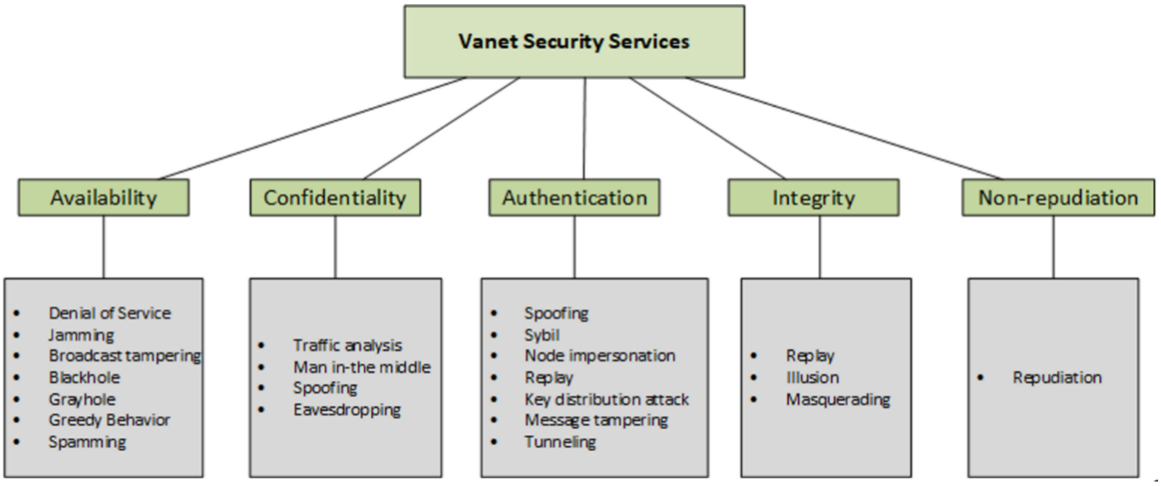

***Figure 2**. Classification of attacks in VANET architectures according to security requirements*

As a result of the examined studies; it has been realized that the subject of machine learning comes to the fore in the new generation security solutions, considering the variable traffic conditions and security criteria in VANET systems. For these reasons, in Section B.3, information about anomaly detection in VANET systems and machine learning algorithms used in IDS systems will be given.

B.3 Machine Learning Algorithms for VANET IDS

Security is an important issue in VANET systems because many attack types compromise the communication of moving vehicles. When the literature studies are examined, many trusted approaches based on IDS (intrusion detection system) have been put forward to against VANETs attacks. But, when the vehicle numbers increase, the data collected by IDSs increases significantly, which causes the detection time to be very long. In addition, the diversity of the collected and analyzed data and the wide range of possible attacks directly increase the variety of machine learning algorithms to be used in IDSs. Therefore, many VANET IDSs in the literature using various machine learning algorithms are intended to improve the detection rate and decrease overheads such as detection time.

SVM algorithm-based anomaly detection, frequently used in the Vanet IDS studies examined, is preferred to separate malicious vehicles in clustered VANETs, detect intrusions and delay communication problems that may occur by detecting attacks. The systems were trained to determine which vehicles exhibited normal behavior on the data collected in the studies, so it was determined whether the vehicle was malicious when new data was shown. When the studies are examined, it has been observed that this supervised machine learning algorithm, which is used in regression and classification in systems with fewer datasets (therefore, less training data), classifies the presence of attacks and anomalies with high accuracy [54, 22, 24, 29, 36, 37]. However, increasing the collaborating vehicles numbers in the systems requires careful tuning of many parameters such as choosing the appropriate kernel in the SVM algorithm, regularization and hyperparameters to fit the dataset, and the expected performance may not be achieved by increasing the complexity. In such cases, it has been observed that methods such as Random Forest that are resistant to complex data, community-based and include more than one independent decision tree are preferred. In VANET IDS studies, in which Random Forest algorithm is used, it has been said that when the training data is large, it performs better than the SVM algorithm and the classification is more accurate [25, 31].

Due to the diversity of the data obtained in VANET IDS studies and the high data density on the network, the performance of the network decreases and it becomes difficult to detect abnormal situations. Therefore, in many studies, it has been seen that various machine learning algorithms are used together or cooperate with existing optimization techniques to detect abnormal situations in IDSs and improve performance in cases such as untrusted connections, and heterogeneous data. It is aimed at detecting malicious attacks of malicious nodes such as packet dropping, resource exhaustion, selective forwarding, and wormhole in IDS studies using the hybrid algorithm examined [17, 26, 27, 28, 34]. In these studies, in addition to the SVM classification algorithm, which is frequently used for anomaly detection, the Dolphin Swarm Optimization algorithm [17], which is used to optimize the detection and accuracy of IDSs, and the C5 Decision Tree Classifier [27] for well-known intrusion detection, to increase the performance of the classifier and to provide high quality It has been observed that optimization or machine learning algorithms such as the K-Means algorithm [28] are highly preferred to create a data set. When IDSs using hybrid algorithms are compared with studies using single algorithms that perform better among existing techniques, it has been seen that the integrated operation of multiple algorithms provides significantly better performance in terms of false positives, detection time, detection rate, etc. parameters during the detection of attacks. In addition, it has been numerically stated that the efficiency of the hybrid IDS models designed in the studies is higher than the IDSs using single algorithms.

In the VANET IDS studies examined, Deep Learning techniques were frequently used during anomaly detection. In these studies, DCNN (Deep convolutional neural network) algorithm was generally preferred. A type of multilayer perceptron, the DCNN algorithm includes a subsampling layer, one or more convolution layers, and one or more fully connected layers such as a standard multilayer neural network. Basically, CNN uses standard neural networks to solve the classification problem but uses other layers to identify information and detect some features [33, 23]. When the studies using DCNN deep learning algorithm in VANET IDS are compared with those using traditional machine learning algorithms such as ANN, LSTM, SVM, decision trees, kNN it has been observed that the obtained FNR and ER parameters are better. In addition, it is known that the wireless and autonomy features of UAVs, which are autonomous vehicles controlled by people with remote controls, cause security vulnerabilities. In the studies on this subject, it has been seen that the DFFNN (deep feed forward neural network) algorithm, which is one of the deep learning models, is used to monitor and analyze the network traffic of UAVs and to detect intruders, giving successful results [60, 61]. Therefore, it has been emphasized in the studies that deep learning algorithms perform better in case of complex, irregular and random attacks on VANETs [33, 23].

The literature shows that the Fuzzy Logic method is combined with various machine learning algorithms to decide whether a sample collected from VANET IDSs represents an anomaly. In a study on intrusion detection in VANET networks [55], the ANFIS model, an intelligent Neuro-Fuzzy technique used to model and control poorly defined and uncertain systems, was used. Thus, it shows successful prediction performance with better accuracy and a lower error rate. Different integrators used in ANFIS models are generally preferred in predictive applications. Thus, the goal is to minimize estimation errors. The ANFIS model is based on input/output data pairs. ANFIS is generally chosen to resolve the problem of constant variability in mobile learning environments in VANET systems and to facilitate the creation of adaptive learning content. Therefore, the ANFIS model can be preferred for unauthorized access detection in VANET IDS. In another study, Fuzzy Logic and Genetic Algorithms were combined for anomaly detection in the network. In this study, Genetic Algorithm extracted information from network flow data to predict network traffic behavior in each time period. The information obtained is used to generate the Digital Signature of the Network Segment. Fuzzy logic scheme is preferred to evaluate whether the obtained samples are abnormal or not [15]. The results obtained from applying the suggested anomaly detection system in a real network traffic flow, achieving 96.53 % accuracy and 0.56 % false positive rate, are said to be more successful than the CNN and SVM algorithms tested in practice. Since the data in VANET systems change very quickly, it has been seen that unsupervised algorithms are used in many studies to reflect the state of these unstable data in real-time and to extract the characteristics of the traffic. For example, this study aims to increase efficiency by modifying the Ant Colony Optimization metaheuristic that optimizes the analysis of multidimensional flow attributes through self-organizing agents and allows timely completion to reduce the impact on large-scale networks [56]. In another study, the state of each vehicle in VANETs was modeled as an HMM (Hidden Markov Model) to filter messages rather than detect messages from vehicles quickly. A Baum-Welch algorithm is first used to generate an HMM for each neighbor tool and its parameters.

Multiple HMMs with their parameters are then used to predict the future states of neighboring vehicles. It is said that the proposed IDS with FM (filter model)- HMM has better performance than conventional methods according to detection rate, detection time and overhead parameters [35]. Rahal et al. [66] suggested a behavior-based multilayer architecture for detecting vehicle botnets in vehicular networks. The suggested architecture comprises two dual-layer modules, the first of which monitors vehicle actions at the network level and the second of which monitors vehicle activities inside the vehicle. Training procedures for these modules were carried out using decision tree techniques. The suggested model has a detection rate of more than 97 % and a false positive rate of less than 0.14 % in testing.

III. RELATED WORKS AND COMPARISON

A. LITERATURE SUMMARIES

This section summarizes some of the recent studies in the literature on machine learning-based VANET architecture. A novel anomaly detection dependent NADS is proposed and tested in [15]. DSNSF generation using a Genetic Algorithm and anomaly identification using a Gaussian Membership function are the two steps of the method. An alert may be activated based on the combined score of the membership degrees, implying that a problem with network traffic can exist. With a high accuracy rate of 96.53 percent and a low false positive rate of 0.56 percent, the system outperforms rigid limits, outlier identification, CkNN, and SVM. In addition, the authors claim that compared to the ACODS solution, the method produced a better overall result. It’s also worth noting that, according to them, this approach is self-contained. Even when up to 40% of vehicles are rogue, the proposed IDS in [16] is still impressive, as its accuracy rate approaches 99.69 percent. The function of traffic flow is determined according to the traffic flow feature of vehicles in the proposed extraction algorithm. The algorithm intends to remove distinct features from vehicle communications in a limited time. It employs an I-GHSOM-based classifier and processes for relabeling and recalculating. In a VANET world, the Dolphin Swarm Optimized IDS is proposed to prevent malicious nodes from executing malevolent attacks like selective routing, packet dropping, resource exhaustion, wormhole, etc. Sharma et al.’s method [17] has an overall detection rate of about 99 percent with node densities ranging from 50 to 300 nodes, with an average false positive rate of 0.87 percent and an average detection time of 44 milliseconds. The system seems to do well regarding efficiency assessment metrics for node densities ranging from 50 to 300 nodes.

A series of specification rules based on the Packet Forwarding Rate (PFR), Packet Drop Rate (PDR), Duplicate Packet Rate and Receive Signal Strength Indicator (RSSI) are used in the proposed IDS system (DPR) in [18]. A CH election algorithm and an IDS Framework based on game theory have also been introduced. The proposed clustering algorithm ensures the IDS framework's stability by creating stable vehicular clusters with improved communications among member vehicles. By studying the efficiency of numerous other classifiers, including the Decision Tree, we hope to boost the DR and reduce the FAR rate. In modern applications like self-driving cars, intelligent intrusion detection systems have become an important security technology. Grey hole and rushing attacks attempt to lose any or all incoming texts, which may significantly affect the lives of passengers, drivers, and cars. Alheeti et al. [19] suggest an adaptive protection scheme for self-driving vehicles’ exterior communication and prevention of grey-hole threats. The system is based on two scenarios created and simulated in NS2 to define normal and abnormal vehicle activity. Based on machine learning approaches a modern method for misbehavior identification is suggested. The model involves four phases: collection, share and review of data in order to extract the representative characteristics and decision making in [20]. The authors’ forthcoming articles will discuss the findings and descriptions of the proposed definition. Yu et al. [21] suggest the development of an open-source Floodlight controller in SDN based on a DDoS attack detection device platform. Suggest a packet-in detection mechanism to substantially minimize the response time to the attack. By experimentation, the author examines the feasibility of the scheme. We are using an algorithm for classifying the samples and developing a detector model to decide whether there is a network attack. Yao et al. [22] discuss using an RSSI-based process to discuss Sybil attacks with power control in VANETs. In PCISAD, a multiplex detection system that combines DTW and CPD and uses periodical detections is developed. Furthermore, an SVM classification system separates Sybil nodes from regular ones.

Shu et al. [23] share intrusion detection method (CIDS) for VANETs is proposed. CIDS allows distributed SDN controllers to learn an effective intrusion detection model collaboratively. The correctness of CIDS in both IID and non-IID cases is shown using robust proofs. Its correctness in both IID and non-IID conditions is illustrated using theoretical research and experimental assessment on a real-world dataset. Deep learning combines generative adversarial networks to provide an effective intrusion prevention approach that can reduce device computing and connectivity load. The issue of detecting malicious vehicles in clustered VANETs was emphasized in the [24]. Cluster participants are assigned to CEAP to collect and monitor evidence about the actions of Multipoint relay (MPR) nodes. And then, using the SVM learning strategy, MPRs are identified as cooperative or malicious. The Gaussian Radial Basis Function kernel marginally outperforms the other functions regarding consistency, false positive rates and attack detection. A massive sharing scheme [25] is proposed to enhance mutual information while reducing collaboration overhead. As compared to the current CIDS model, the general performance according to F1 score was 97 percent with a 4% false positive rate. According to the scientists, MA-CIDS outperforms all other current VANET models. According to the researchers, the collective IDS paradigm would be explored using controlled and unsupervised machine learning techniques.

Ghaleb et al. introduces a community hybrid context-sensitive misbehavior recognition model [26]. EHCA-MDS adheres to the principles of crowd intelligence and the strength of plurality. These classifiers collaborate to detect various forms of misbehaving cars. According to the researchers, the model has shown robustness and adaptability even in heterogeneous noise conditions and unreliable connectivity. The efficiency of the proposed model is 10% better than the hybrid model and 37% higher than the data-centric model. In [27], the C5 classifiers generate an intruder signature, which can generate a law pattern more quickly and detect intrusions with fewer signatures. It is also shown that our Hybrid IDS outperforms current techniques. As compared to single algorithms, integrating multiple algorithms yielded significantly better performance. Compared to other machine learning techniques, an ensemble of the C5 classifier (signature) and one-class SVM (anomaly) will result in a higher detection rate. The aim of [28] is to create a model that can cope with real-world intrusion detection problems in data processing. It suggests a multi-level hybrid intrusion detection model that uses support vector machines and extreme learning machines. The proposed model is highly effective at detecting threats, and its accuracy (95.75 percent) is the highest. Network attacks are more likely in a VANET system with ad hoc routing and highly complex topology. It is discovered by allowing each vehicle to track its next hop in the packet routing path for three crucial parameters. A basic TVT ensures that legitimate vehicles are rewarded with higher trust values, while in this study, misbehaving nodes that disrupt packet transmission are penalized [29].

Because of the complex nature of their clients and the amount of information exchanged between them and their respective networks, VANETs are especially difficult to secure. As a potential alternative, a spline-based intrusion detection method has been pioneered [30]. This knot flow classification approach (KFC) facilitates efficient intrusion detection by integrating clustering with spline-based general linear model classification. The proposed framework [31] uses Spark to accelerate data collection and HDFS to store large amounts of suspected attacks. The Random Forest classification algorithm is used in the traffic detection module. The suggested detection algorithm achieved an accuracy rate of 99.95 % and 98.75 %, respectively, with a false warning rate of 0.05% and 1.08%. Idhammed et al. [32] present an online sequential semi-monitored DDoS identification method, based on network estimates, co-clusters, data gain ratios, and extra-trees. Several trials were conducted to test the alternative solution using three public datasets. NSL-KDD, UNB ISCX 12, and UNSW-NB15 achieve accuracy of 98.23 percent, 99.88 percent, and 93.71 percent, respectively.

The focus of [33] is to develop an intrusion detection model that learns sequential network traffic patterns in vehicles and detects message injection attacks. The proposed IDS was developed using the Inception-ResNet structure initially intended for large-scale imaging. It demonstrated that the intrude detection of attacks by the DCNN-based IDS is accurate and effective, including DoS, RPM spoofing, gear spoofing and regular traffic fuzzy attacks. The proposed DCNN model showed a large increase in efficiency relative to other algorithms for the fuzzy attack dataset. Adhikary et al. [34] propose a hybrid algorithm for detecting DDoS attacks in Vanet based on the AnovaDot and RBFDot SVM kernel methods. Features such as colliding, packet drop, jitter etc. have been used for simulating the network connectivity situation in real-time in this hybrid algorithm. The test results show that the model built on this algorithm is greater than the SVM kernel-based models. The proposed FM-HMM [35] is efficient for almost all types of IDSs, implying that it can boost general IDS efficiency. According to simulations, network message latency can be reduced by reducing message size and improving processing speed. And though up to 40 % of cars are rogue, the efficiency of the proposed IDS is still amazing. In the VANET environment, Shams et al. [36] use a poly nucleic kernel of degree 2 to train SVM IDM for packet falling and delaying DoS attack detection. With high PR and RC, the device will classify the presence or absence of DoS attackers with an accuracy of 99 percent. The monitoring module has a low FAR to discourage false alerts from affecting network availability. The hybrid IDS model of Alsarhan et al.[37] is specifically concerned with avoiding cyber-security attacks in VANET. Integrating several pieces of data, a Dempster-Shafer hypothesis is used to quantify the probability of attacks. Based on the simulation findings, we infer that the optimal strategy is obtained in IDS by combining various evidence and understanding. The proposal is based on the convergence of three approaches: rule-based filtering, event-specific trust, and prior information. Kerrache et al. [38] present a process design for effective trust establishment in VANETs. We concentrate on how nodes should stop sending legal messages over untrusted or unreliable connections. We implemented the concept of companions, which incorporates the concepts of relationship stability and neighbor confidence value. Also, in dynamic situations such as DoS or DDoS attacks, the scheme will ensure a high identification ratio of dishonest nodes in the network. In [37], a support vector machine (SVM) is used to detect intrusions in a VANET. Smart vehicles are a promising technology for intelligent transportation systems (ITS) in smart cities. However, the free wireless medium makes smart vehicle connectivity more vulnerable to cyber-security threats. Regarding classification precision, the paper compared the efficiency of three ML algorithms on the SVM. Compared to other algorithms, it was discovered that the GA algorithm outperformed its counterparts in classification accuracy.

A Private-Collaborative and secure Intrusion Detection System (SP-CIDS) is proposed for detecting network attacks and mitigating security problems [40]. The framework is used for the Alternating Direction Method of Multipliers with the Distributed Machine Learning (DML) algorithm. V2V coordination during the learning process may increase the IDS’s scalability and accuracy. But major data privacy issues could emerge due to such a partnership. The simulation results show that a private ensemble classifier secures the training data with DP and obtains 96.94 % accuracy. Nadarajan et al. [41] suggest a messy routing scheme based on QoS-aware for VANET end-to-end communication. With the newly boosted LSTM algorithm, the proposed SCARP routing system is extremely predictable. Apart from in the presence of a higher degree of attack, the efficiency of the proposed algorithm is stable and strong. This paper provides new perspectives on the way to incorporate the messy transmission. It underlines the implementation of various hazard models for validation. Diaz et al. [42] offer prediction models as the foundation for future study. These models endorse intrusion detection mechanisms dependent on abnormalities. One benefit of the paradigm is that neural networks are a means of learning. ANFIS handles more difficult parameters resulting by the reduction of variables by the APC. This shows that our protection index suddenly varies with usual network activity and its consequences when a vehicle is attacked.

B. THE COMPARISON OF RECENT STUDIES IN THE LITERATURE

Recent studies are compared on certain criteria. In the comparison, the publication year, attack types, machine learning model used, simulation tool, feature extraction, the machine learning algorithm used by the system, the dataset used in the study, and the performance criteria obtained are DR, FPR, Precision, Recall, and F-Measure values are given. Although there are many studies on this subject in the literature, studies belonging to the last years were selected while making the selection. Figure 3 shows the distribution of the publications examined in the article by year.

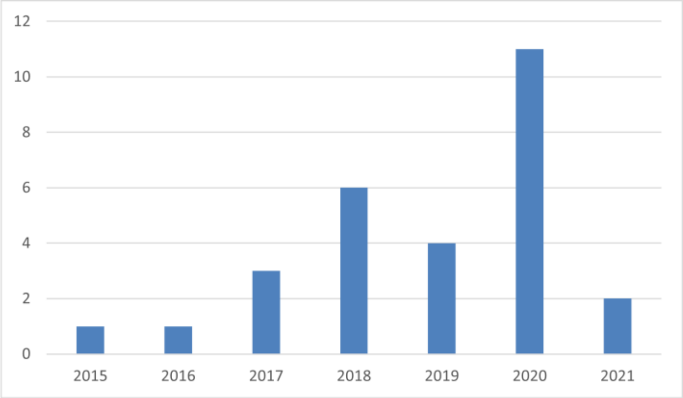

Figure 3. *The distribution of publications by year in the studies*

When the studies in the literature are examined, many different attacks have been carried out on VANET networks. These attack types threaten different security services. Explanations about these attack types are given in Section 2.2.2. Figure 4 shows the numerical distribution of attack types in the articles reviewed. According to this graph, it is seen that more studies are done on DOS and DDoS attacks than other attack types. After DoS and DDoS attacks, it is seen that most studies on normal and attack, that is, binary classification, are carried out. Notably, the number of studies on some attack types is quite low compared to others.

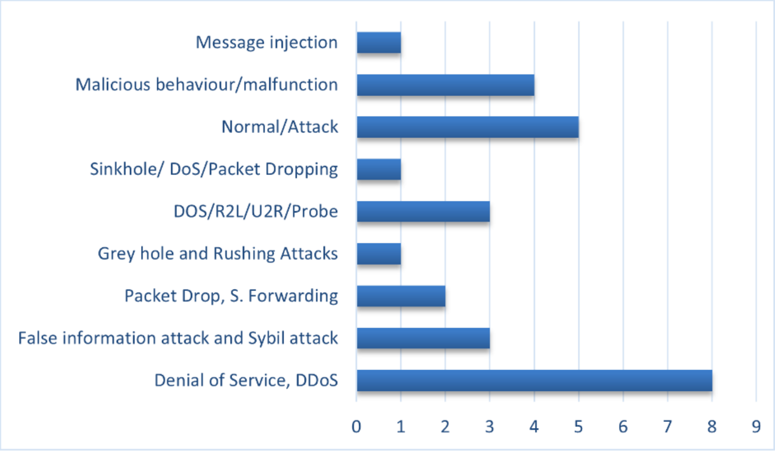

Figure 4. *The distribution of attack types used in the studies.*

There are different learning models used in Machine Learning studies. At the beginning of these models are supervised and unsupervised approaches. In these approaches, many different algorithms are used [43]. The machine learning based VANET IDS studies in the literature mostly use the supervised approach. Figure 5 shows the distribution of the articles examined according to the machine learning model. When Figure 5 is examined, it has been determined that the studies conducted using the supervised/classification technique are considerably higher than the others.

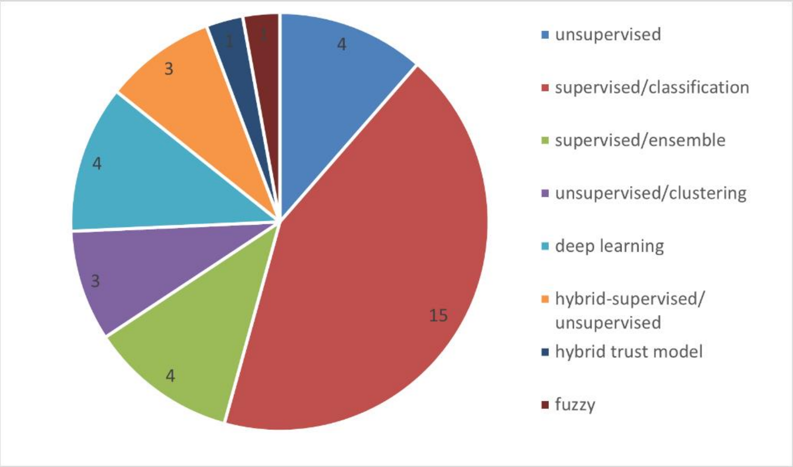

Figure 5. *The distribution of the learning model used in studies*

Many machine learning algorithms have been developed using different approaches in the literature so far. In Section 2.3, an evaluation of machine learning algorithms, which are frequently used in VANET networks, is presented. Figure 6 shows the distribution of machine learning algorithms used in the articles examined in the study. According to the distribution results, it was determined that hybrid methods were preferred more than other methods, ensemble techniques were widely used in these studies, and the SVM algorithm was also used more than other methods. However, it has been used in many articles that used machine learning algorithms.

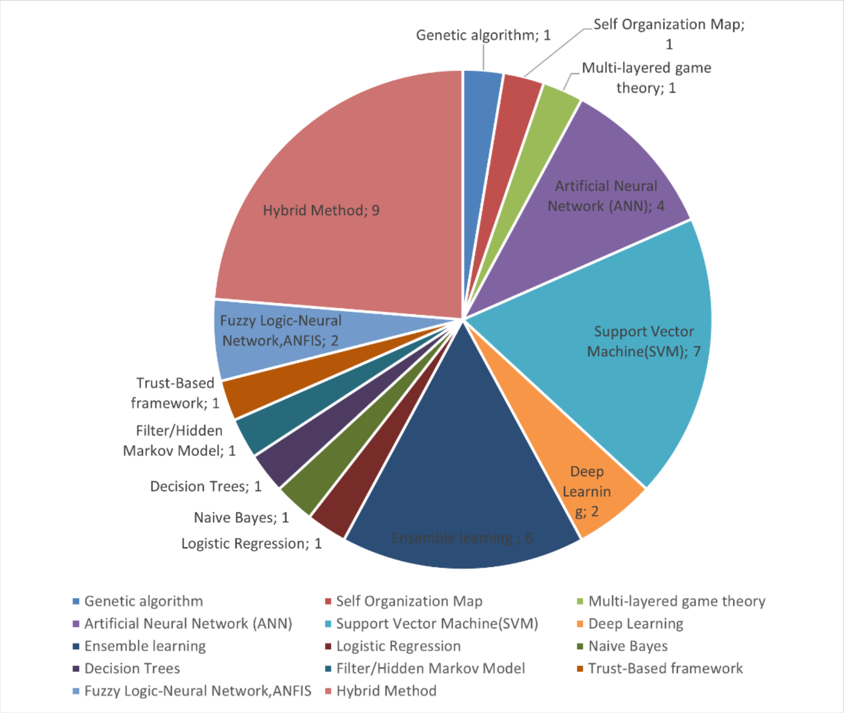

Figure 6. *Distribution of studies in the literature according to machine learning algorithms*

There are many different simulation tools used in VANET network simulations. Operations are carried out by using these simulation tools individually or together in studies. Among these simulation tools, NS-2 [44], NS-3 [45], SUMO [46], VanetMobiSim [47] are among the most commonly used simulators in the literature. Figure 7 shows the distribution of simulation tools used in the studies. It is seen that the NS-2/SUMO simulation tools are mostly used together in the studies and are the most preferred simulation tools among the studies examined, and the NS-2 simulator is also the most preferred simulator as a singular one in the studies. Personal simulation applications written in Phyton/R programming language are also common, and different simulators, which are widely used in the literature, have been preferred in other studies.

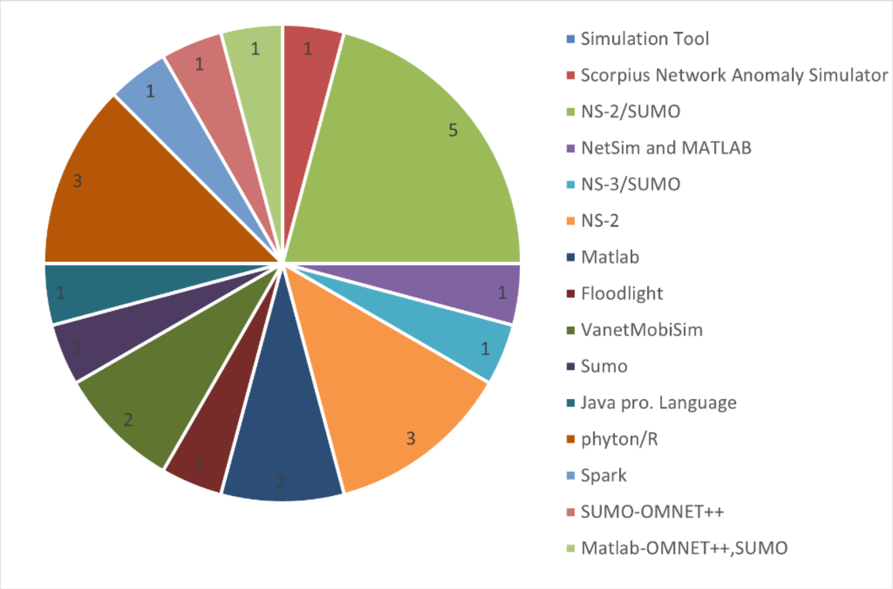

Figure 7. *The distribution of simulation tools used in studies.*

Generally produced data sets are used to test the performance of VANET IDS systems, but in some of the studies, ready data sets are used. Among these datasets, KDD-CUP99 [48], NSL-KDD [49], UNSW-NB15 [50], NGSIM [51] datasets are preferred in studies. Figure 8 shows the dataset distribution used for the training and testing of the preferred system in the studies. It is seen that the authors mostly create datasets by obtaining traffic through their own test environments. Datasets produced in almost half of the studies examined were used. However, NSL-KDD data set is widely preferred in studies. It has been determined that different and newer data sets in the literature are used in the studies.

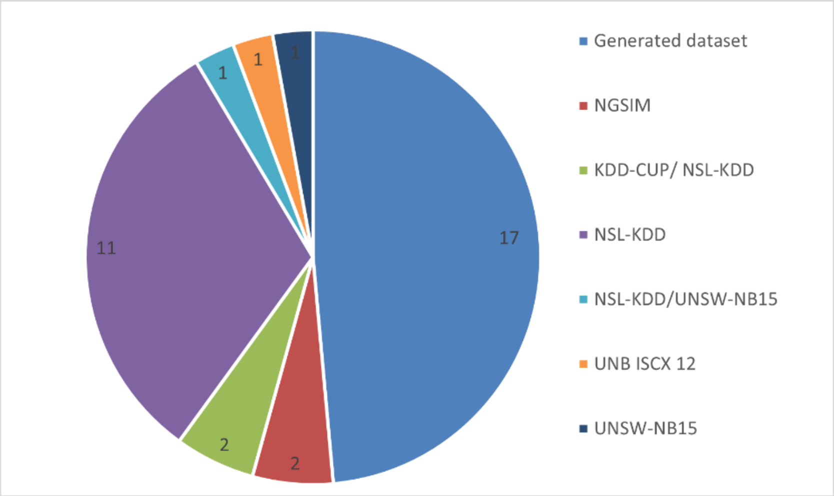

Figure 8. *The distribution of datasets used in studies*

There are many different parameters used in the performance of intrusion detection systems. Among these parameters, Accuracy (ACC) and False Positive Rate (FPR) values are critical for the performance of the system, and many studies have been evaluated using these two criteria [52, 53]. The complexity matrix is given in Table 2. The equations used to calculate the Accuracy, True Positive Rate, and false positive rate values are shown in Equations 1,2 and 3, respectively.

Table 2. *The Confusion Matrix*

2*		**True Class**	
		Positive	**Negative**
2*Predicted Class	Positive	TP	FP
	Negative	FN	TN

$$\mathrm{ACC} = \frac{\mathrm{TP+TN}}{\mathrm{P+N}} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \tag{1}$$

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{P}} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} = 1 - \mathrm{FNR} \tag{2}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{N}} = \frac{\mathrm{FP}}{\mathrm{FP+TN}} = 1 - \mathrm{TNR} \tag{3}$$

Figure 9. compares ACC rates obtained in the studies reviewed. According to the accuracy results obtained in the studies, it was determined that Random Forest, SOM, FFFN, SVM, and ANN algorithms were used in the studies where the highest ACC value was obtained, studies were carried out with NS2-SUMO simulation tools, and studies were mostly carried out on generated datasets with the Unsupervised / Classification model. When we look at the lowest results obtained, it was determined that SVM, RF and Bayes classifier were used, more ready-made datasets (NSL-KDD and UNSW-NB15) were preferred, and the supervised/classification model was chosen. As a result, it can be concluded that the established model and the preferred dataset and simulator influence the performance, since the same machine learning algorithms are also included in the highest and lowest results.

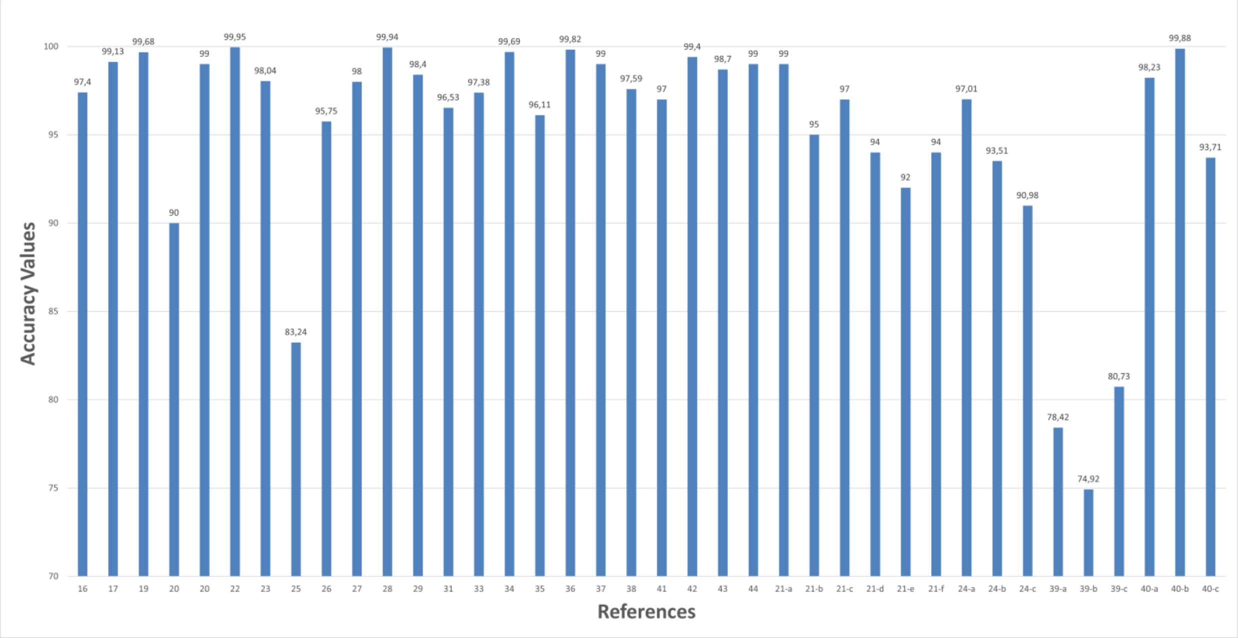

Figure 9. *The comparison table of Accuracy values according to reference numbers*

Another important criterion in the evaluation of intrusion detection systems is FPR. A low FPR value is an important indicator for the system's performance. In the articles reviewed in Figure 10, it was concluded that ensemble learning algorithms, RF, ANN, CNN and SVM algorithms were used as machine learning algorithms, the supervised/ensemble model was preferred, and simulation processes were performed on SUMO and spark in the articles with the lowest FPR value. In the articles with the highest results regarding the FPR criteria, it was observed that SVM and game theory algorithms were preferred together with the supervised/classifying model, and studies were carried out on the generated datasets in NS-2/NS-3/SUMO simulation environments.

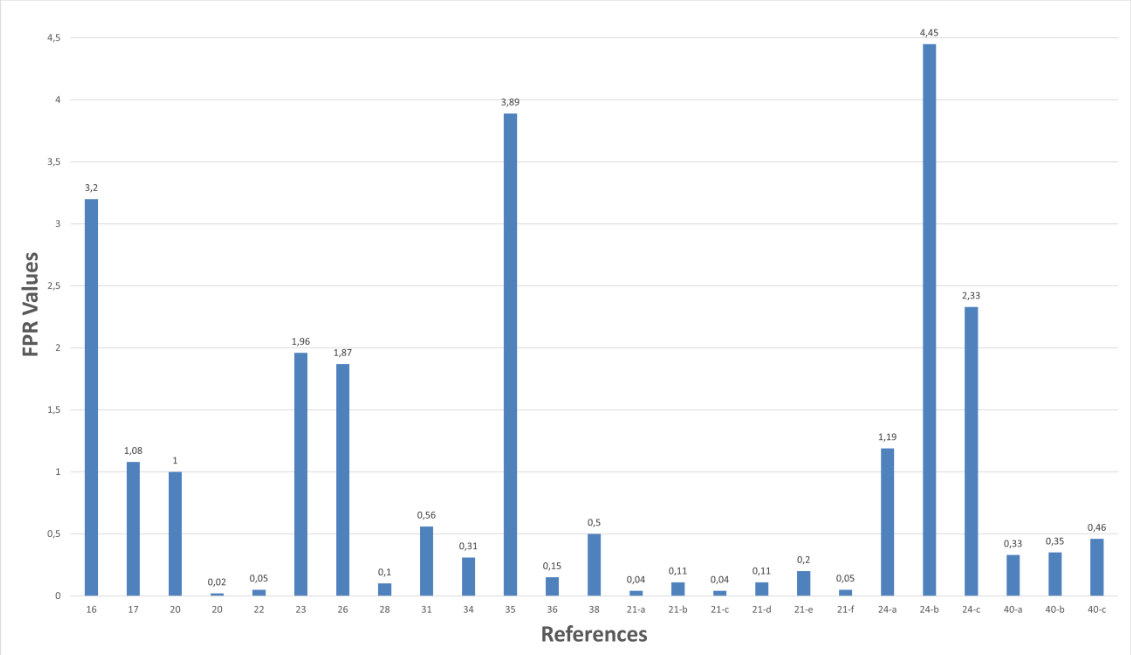

Figure 10. *The comparison table of FPR values according to reference numbers*

IV. CONCLUSION

In this article, a comprehensive review of machine learning studies used for attack detection in VANET systems has been carried out. Recent studies in the literature were selected and evaluated according to different criteria. In the study, information about VANET architectures is first given, and then security requirements are explained. Then, summaries of the studies made in the articles selected from the studies in the literature were presented and a comparison table was created.

The distribution of the examined studies by years is given and it is stated that the studies in this field have increased in recent years. Then, a classification was made according to attack types, and it can be said that DOS/DDoS attacks are more than other attack types; as a result, they are the most dangerous attack types on VANET systems. Another inference in the examinations is that the supervised/classification model is preferred as the learning model used. When the results obtained according to machine learning algorithms are evaluated, it is seen that hybrid methods, ensemble techniques, and SVM algorithms are more commonly used in intrusion detection systems. It has been observed that the combined use of NS-2/SUMO simulation tools is common in the studies and that the authors use datasets obtained from the traffic they generate in the simulation environment they have prepared. However, there are studies on different simulators and datasets. According to the performance criteria results, it has been shown that high-performance studies were performed by obtaining high results for the ACC value and very low results for the FPR value. As a result, it can be said that VANET systems are exposed to many different attacks due to their architecture, and high-performance results are obtained by using different machine learning algorithms on different platforms to detect and prevent these attacks, and these studies are increasing with new approaches.

V. REFERENCES

[1] W. H. Organization et al., “Global status report on road safety 2018: summary,” World Health Organization, Tech. Rep., 2018.

[2] M. Alam, J. Ferreira, and J. Fonseca, “Introduction to intelligent transportation systems,” in *Intelligent transportation systems*. *Springer*, 2016, pp. 1–17.

[3] C. T. Barba, M. A. Mateos, P. R. Soto, A. M. Mezher, and M. A. Igartua, “Smart city for vanets using warning messages, traffic statistics and intelligent traffic lights,” in 2012 IEEE intelligent vehicles symposium. IEEE, 2012, pp. 902–907.

[4] P. K. Singh, S. K. Nandi, and S. Nandi, “A tutorial survey on vehicular communication state of the art, and future research directions,” *Vehicular Communications*, vol. 18, p. 100164, 2019.

[5] L. Sumi and V. Ranga, “An iot-vanet-based traffic management system for emergency vehicles in a smart city,” in *Recent Findings in Intelligent Computing Techniques*. Springer, 2018, pp. 23–31.

[6] S. Abdelatif, M. Derdour, N. Ghoualmi-Zine, and B. Marzak, “Vanet: A novel service for predicting and disseminating vehicle traffic information,” *International Journal of Communication Systems*, vol. 33, no. 6, p. e4288, 2020.

[7] A. M. De Souza, C. A. Brennand, R. S. Yokoyama, E. A. Donato, E. R. Madeira, and L. A. Villas, “Traffic management systems: A classification, review, challenges, and future perspectives,” *International Journal of Distributed Sen-sor Networks*, vol. 13, no. 4, p. 1550147716683612, 2017.

[8] H. Hasrouny, A. E. Samhat, C. Bassil, and A. Laouiti, “Vanet security challenges and solutions: A survey,” *Vehicular Communications*, vol. 7, pp. 7–20, 2017.

[9] M. S. Sheikh and J. Liang, “A comprehensive survey on vanet security services in traffic management system,” *Wireless Communications and Mobile Computing*, vol. 2019, 2019.

[10] M. B. Mansour, C. Salama, H. K. Mohamed, and S. A. Hammad, “Vanet security and privacy-an overview,” *International Journal of Network Security Its Applications (IJNSA)* Vol, vol. 10, 2018.

[11] R. Abassi, “Vanet security and forensics: Challenges and opportunities,” *Wiley Interdisciplinary Reviews: Forensic Science*, vol. 1, no. 2, p. e1324, 2019.

[12] R. Hussain, F. Hussain, and S. Zeadally, “Integration of vanet and 5g security: A review of design and implementation issues,” *Future Generation Computer Systems*, vol. 101, pp. 843–864, 2019.

[13] A. N. Upadhyaya and J. Shah, “Attacks on vanet security,” *Int J Comp Eng Tech*, vol. 9, no. 1, pp. 8–19, 2018.

[14] M. A. Hezam, A. Junaid, A. Syed, M. Nazri, M. Warip, K. N. Fazira, K. Azir, and R. Nurul Hidayah, “Classification of security attacks in vanet: A review of requirements and perspectives,” 2018.

[15] D. Manivannan, S. Moni, and S. Zeadally, “Secure authentication and privacy-preserving techniques in vehicular ad-hoc networks (vanets),” *Vehicular Communications*, vol. 25, p. 100247, 02 2020.

[16] T. Nandy, M. Y. I. Idris, R. M. Noor, A. W. A. Wahab, S. Bhattacharyya, R. Kolandaisamy, and M. Yahuza, “A secure, privacy-preserving, and lightweight authentication scheme for vanets,” *IEEE Sensors Journal*, vol. 21, no. 18, pp. 20 998– 21 011, 2021.

[17] M. A. Al-Shareeda, M. Anbar, I. H. Hasbullah, and S. Manickam, “Survey of authentication and privacy schemes in vehicular ad hoc networks,” *IEEE Sensors Journal*, vol. 21, no. 2, pp. 2422–2433, 2021.

[18] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, “High-dimensional and large-scale anomaly detection using a linear one-class svm with deep learning,” *Pattern Recognition*, vol. 58, pp. 121–134, 2016.

[19] Y. Yao, B. Xiao, G. Yang, Y. Hu, L. Wang, and X. Zhou, “Power control identification: A novel sybil attack detection scheme in vanets using rssi,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2588–2602, 2019.

[20] O. A. Wahab, A. Mourad, H. Otrok, and J. Bentahar, “Ceap: Svm-based intelligent detection model for clustered vehicular ad hoc networks,” *Expert Systems with Applications*, vol. 50, pp. 40–54, 2016.

[21] E. A. Shams, A. Rizaner, and A. H. Ulusoy, “Trust aware support vector machine intrusion detection and prevention system in vehicular ad hoc networks,” *Computers Security*, vol. 78, pp. 245–254, 2018.

[22] E. A. Shams, A. H. Ulusoy, and A. Rizaner, “Performance analysis and comparison of anomaly-based intrusion detection in vehicular ad hoc networks.” *Radioengineering*, vol. 29, no. 4, 2020.

[23] A. Alsarhan, A.-R. Al-Ghuwairi, I. T. Almalkawi, M. Alauthman, and A. Al-Dubai, “Machine learning-driven optimization for intrusion detection in smart vehicular networks,” *Wireless Personal Communications*, vol. 117, no. 4, pp. 3129–3152, 2021.

[24] F. A Ghaleb, F. Saeed, M. Al-Sarem, B. Ali Saleh Al-rimy, W. Boulila, A. Eljialy, K. Aloufi, and M. Alazab, "Misbehavior-aware on-demand collaborative intrusion detection system using distributed ensemble learning for vanet," *Electronics*, vol. 9, no. 9, p. 1411, 2020.

[25] Y. Gao, H. Wu, B. Song, Y. Jin, X. Luo, and X. Zeng, "A distributed network intrusion detection system for distributed denial of service attacks in vehicular ad hoc network," *IEEE Access*, vol. 7, pp. 154 560–154 571, 2019.

[26] S. Sharma and A. Kaul, "Hybrid fuzzy multi-criteria decision making based multicluster head dolphin swarm optimized ids for vanet," *Vehicular Communications*, vol. 12, pp. 23–38, 2018.

[27] F. A. Ghaleb, M. A. Maarof, A. Zainal, B. A. S. Al-rimy, A. Alsaeedi, and W. Boulila, "Ensemble-based hybrid context-aware misbehavior detection model for vehicular ad hoc network," *Remote Sensing*, vol. 11, no. 23, p. 2852, 2019.

[28] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, "Hybrid intrusion detection system based on the stacking ensemble of c5 decision tree classifier and one class support vector machine," *Electronics*, vol. 9, no. 1, p. 173, 2020.

[29] W. L. Al-Yaseen, Z. A. Othman, and M. Z. A. Nazri, "Multi-level hybrid support vector machine and extreme learning machine based on modified k-means for intrusion detection system," *Expert Systems with Applications*, vol. 67, pp. 296–303, 2017.

[30] K. Adhikary, S. Bhushan, S. Kumar, and K. Dutta, "Hybrid algorithm to detect ddos attacks in vanets," *Wireless Personal Communications*, vol. 114, no. 4, pp. 3613–3634, 2020.

[31] H. M. Song, J. Woo, and H. K. Kim, "In-vehicle network intrusion detection using deep convolutional neural network," *Vehicular Communications*, vol. 21, p. 100198, 2020.

[32] J. Shu, L. Zhou, W. Zhang, X. Du, and M. Guizani, "Collaborative intrusion detection for vanets: a deep learning-based distributed sdn approach,"*IEEE Transactions on Intelligent Transportation Systems*, 2020.

[33] V. N and P. Ganapathi, Traffic Analysis of UAV Networks Using Enhanced Deep Feed Forward Neural Networks (EDFFNN), 01 2020, pp. 219–244.

[34] H. Bangui and B. Buhnova, "Recent advances in machine-learning driven intrusion detection in transportation: Survey," *Procedia Computer Science*, vol. 184, pp. 877–886, 01 2021.

[35] B. A. Bensaber, C. G. P. Diaz, and Y. Lahrouni, "Design and modeling an adaptive neuro-fuzzy inference system (anfis) for the prediction of a security index in vanet," *Journal of Computational Science*, vol. 47, p. 101234, 2020.

[36] A. H. Hamamoto, L. F. Carvalho, L. D. H. Sampaio, T. Abr˜ao, and M. L. Proen¸ca Jr, "Network anomaly detection system using genetic algorithm and fuzzy logic," *Expert Systems with Applications*, vol. 92, pp. 390–402, 2018.

[37] L. F. Carvalho, S. Barbon Jr, L. de Souza Mendes, and M. L. Proenca Jr, "Unsupervised learning clustering and self-organized agents applied to help network management," *Expert Systems with Applications*, vol. 54, pp. 29–47, 2016.

[38] J. Liang, M. Ma, M. Sadiq, and K.-H. Yeung, "A filter model for intrusion detection system in vehicle ad hoc networks: A hidden Markov methodology," *Knowledge-Based Systems*, vol. 163, pp. 611–623, 2019.

[39] R. Rabah, A. K. Abdelaziz, G.-Z. Nacira, C. Yacine, and M. Ghamri-Doudane, “Antibotv: A multilevel behaviour-based framework for botnets detection in vehicular networks,” *Journal of Network and Systems Management*, vol. 30, pp. 1–15, 03 2022.

[40] J. Liang, J. Chen, Y. Zhu, and R. Yu, “A novel intrusion detection system for vehicular ad hoc networks (vanets) based on differences of traffic flow and position,” *Applied Soft Computing*, vol. 75, pp. 712–727, 2019.

[41] B. Subba, S. Biswas, and S. Karmakar, “A game theory based multi layered intrusion detection framework for vanet,” *Future Generation Computer Systems*, vol. 82, pp. 12–28, 2018.

[42] K. M. A. Alheeti, A. Gruebler, and K. D. McDonald-Maier, “On the detection of grey hole and rushing attacks in self-driving vehicular networks,” in 2015 7th Computer science and electronic engineering conference (CEEC). IEEE, 2015, pp. 231–236.

[43] F. A. Ghaleb, A. Zainal, M. A. Rassam, and F. Mohammed, “An effective misbehavior detection model using artificial neural network for vehicular ad hoc network applications,” in 2017 IEEE Conference on Application, Information and Network Security (AINS). IEEE, 2017, pp. 13–18.

[44] Y. Yu, L. Guo, Y. Liu, J. Zheng, and Y. Zong, “An efficient sdn-based ddos attack detection and rapid response platform in vehicular networks,” *IEEE access*, vol. 6, pp. 44 570–44 579, 2018.

[45] D. A. Schmidt, M. S. Khan, and B. T. Bennett, “Spline-based intrusion detection for vanet utilizing knot flow classification,” *Internet Technology Letters*, vol. 3, no. 3, p. e155, 2020.

[46] M. Idhammad, K. Afdel, and M. Belouch, “Semi-supervised machine learning approach for ddos detection,” *Applied Intelligence*, vol. 48, no. 10, pp. 3193–3208, 2018.

[47] C. A. Kerrache, N. Lagraa, C. T. Calafate, and A. Lakas, “Tfdd: A trust-based framework for reliable data delivery and dos defense in vanets,” *Vehicular Communications*, vol. 9, pp. 254–267, 2017.

[48] G. Raja, S. Anbalagan, G. Vijayaraghavan, S. Theerthagiri, S. V. Suryanarayan, and X.-W. Wu, “Sp-cids: Secure and private collaborative ids for vanets,” *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[49] J. Nadarajan and J. Kaliyaperumal, “Qos aware and secured routing algorithm using machine intelligence in next generation vanet,” *International Journal of System Assurance Engineering and Management*, pp. 1–12, 2021.

[50] C. G. P. Diaz, B. A. Bensaber, and Y. Lahrouni, “Design and modeling an adaptive neuro-diffuse system (anfis) for the prediction of a security index in vanet,” in 2019 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2019, pp.1040–1045.

[51] A. Alsarhan, M. Alauthman, E. Alshdaifat, A.-R. Al-Ghuwairi, and A. Al-Dubai, “Machine learning-driven optimization for svm-based intrusion detection system in vehicular ad hoc networks,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–10, 2021.

[52] H. Bangui, M. Ge, and B. Buhnova, “A hybrid machine learning model for intrusion detection in vanet,” *Computing*, 08 2021.

[53] J. Liang, Q. Lin, J. Chen, and Y. Zhu, “A filter model based on hidden generalized mixture transition distribution model for intrusion detection system in vehicle ad hoc networks,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, pp. 2707–2722, 2020.

[54] A. Gad, A. Nashat, and T. Barkat, “Intrusion detection system using machine learning for vehicular ad hoc networks based on ton-iot dataset,” *IEEE Access*, vol. 9, pp. 142 206–142 217, 10 2021.

[55] S. Ercan, M. Ayaida, and N. Messai, “Misbehavior detection for position falsification attacks in vanets using machine learning,” *IEEE Access*, vol. PP, pp. 1–1, 12 2021. [56] T. Jo, Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning. Springer Nature, 2021.

[57] T. Issariyakul and E. Hossain, “Introduction to network simulator 2 (ns2),” in Introduction to network simulator NS2. *Springer*, 2009, pp. 1–18.

[58] G. F. Riley and T. R. Henderson, “The ns-3 network simulator,” in Modeling and tools for network simulation. *Springer*, 2010, pp. 15–34.

[59] P. Fernandes and U. Nunes, “Platooning of autonomous vehicles with intervehicle communications in sumo traffic simulator,” in 13th International IEEE Conference on Intelligent Transportation Systems. IEEE, 2010, pp. 1313–1318.

[60] J. Harri, F. Filali, C. Bonnet, and M. Fiore, “Vanetmobisim: generating realistic mobility patterns for vanets,” in Proceedings of the 3rd international workshop on Vehicular ad hoc networks, 2006, pp. 96–97.

[61] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “A detailed analysis of the kdd cup 99 data set,” in 2009 IEEE symposium on computational intelligence for security and defense applications. IEEE, 2009, pp. 1–6.

[62] L. Dhanabal and S. Shantharajah, “A study on nsl-kdd dataset for intrusion detection system based on classification algorithms,” *International journal of advanced research in computer and communication engineering*, vol. 4, no. 6, pp. 446–452, 2015.

[63] S. Meftah, T. Rachidi, and N. Assem, “Network based intrusion detection using the unsw-nb15 dataset,” *International Journal of Computing and Digital Systems*, vol. 8, no. 5, pp. 478–487, 2019.

[64] B. Coifman and L. Li, “A critical evaluation of the next generation simulation (ngsim) vehicle trajectory dataset,” *Transportation Research Part B: Methodological*, vol. 105, pp. 362–377, 2017.

[65] A. Tharwat, “Classification assessment methods,” *Applied Computing and Informatics*, 2020.

[66] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, “A deep learning approach for network intrusion detection system,” *Eai Endorsed Transactions on Security and Safety*, vol. 3, no. 9, p. e2, 2016.