An Edge-enabled Virtual Honeypot Based Intrusion Detection System for Vehicle-to-Everything (V2X) Security using Machine Learning

S.Thangam, S.Sibi Chakkaravarthy

***Abstract*—Securing vehicle-to-everything (V2X) communications is essential as intelligent transportation system integration progresses to guarantee the dependability and safety of connected vehicles. Our study presents a novel approach aimed at strengthening the security of vehicles in V2X networks. The proposed system utilizes the virtual honeypots technique, referred to as PotRSU, within roadside units (RSU) to gather data from heterogeneous sources. The malicious entities that are drawn from all incoming traffic are recorded by the PotRSU. We utilized machine learning algorithms to effectively identify intrusion. The analysis and experimentation conducted on the proposed system exhibit 99.01% accuracy in identifying malicious nodes.**

***Index Terms*—Vehicular networks, V2X security, Intrusion detection system (IDS), RSU, Honeypot.**

I. Introduction

THE advent of machine learning (ML) and the widespread use of interconnected technologies have brought about a new era in vehicular communication, known as V2X [1]. V2X networks revolutionize the automobile industry by enabling vehicles to connect with each other (V2V) [2], infrastructure (V2I) [3], networks (V2N) [4], and pedestrians (V2P) [5], [6]. The primary objective of this extensive communication architecture is to improve road safety, optimize the traffic flow of vehicles [7], and provide a wide range of new applications and services. V2X communication is based on transmitting critical information between vehicles and their environment (Figure 1). This bidirectional data exchange provides information on vehicle location, speed, and trajectory, therefore enabling a system whereby vehicles learn about one another and their surroundings [8]. This level of connectivity serves as the infrastructure for sophisticated driver assistance systems, autonomous vehicles, and other smart city applications. The V2X ecosystem has a resilient architecture with onboard units integrated within vehicles [9] and roadside units strategically placed alongside routes [10]. These components collaborate to establish an interconnected network that enables smooth communication. Furthermore, cloud-based services facilitate the storage, analysis, and retrieval of vast data produced by connected vehicles [11].

Manuscript received January 21, 2024; revised July 22, 2024.

S.Thangam is a Ph.D. candidate at the School of Computer Science and Engineering, VIT-AP University, Andhra Pradesh, India. (e-mail: thangam.21phd7031@vitap.ac.in).

S.Sibi Chakkaravarthy is an Associate professor at the School of Computer Science and Engineering, Center of Excellence in Artificial Intelligence and Robotics (AIR), VIT-AP University, Andhra Pradesh, India. (e-mail: sb.sibi@gmail.com).

Though V2X communication holds greater potential for transformation, the dynamic nature of these networks brings security issues. The possibilities of unauthorized access, challenges to data integrity, and potential misbehavior of vehicles are substantial. Securing V2X networks involves safeguarding confidential data and the well-being of the physical environment and the individuals who reside within it [12].

With the continuous evolution of the threat landscape, supplementary security measures become indispensable. This is where cutting-edge technologies, such as blockchain [13] and artificial intelligence (AI) [14], come into action. Honeypots, commonly employed in the field of cybersecurity, are currently being utilized in the realm of V2X security [15]. Existing studies (Table I) are not used to identify recent threats created by attackers. Therefore, we propose a system with a dynamic and adaptive component - honeypot and ML [16]. Real-time threat detection is made possible by this predictive capability, facilitating a proactive response to potential security incidents. A virtual honeypot [17], PotRSU, is implemented in the proposed system, which generates controlled environments to attract potential threats. This PotRSU functions as an information-gathering system,

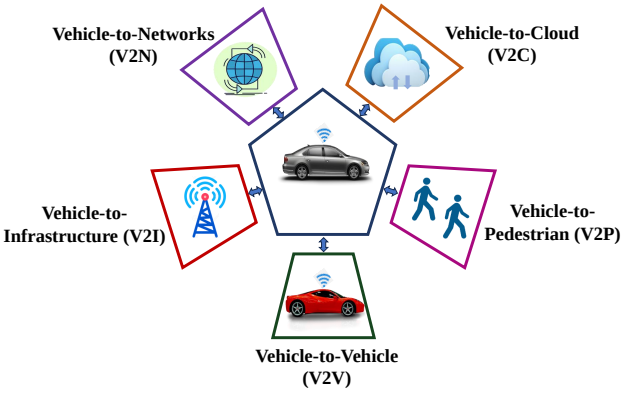


Fig. 1: General structure of V2X.

enabling the detection of malicious patterns and behaviors as well as differentiating anomalous activities that are not overtly malicious. Cloud-based threat intelligence further ensures that the knowledge acquired is not limited to a specific area. The collaborative nature of the system ensures that insights gained from virtual honeypots are shared with real roadside units and cloud-based services. The implementation of bidirectional communication significantly improves the

TABLE I: State-of-the-art of the reviews.

Reference No.	Nature of the Work	Techniques Used	Datasets Used	Advantages	Limitations
[18]	To get the high intrusion detection rate with minimal time and energy consumption	Machine learning	NSL-KDD	Hash chain is used to avoid the intrusions	This model may not be suitable for the new attacks and large datasets
[19]	To classify the normal packets from the malicious packets	Deep learning	KDD-99 and CICIDS 2018	Easy to identify the different types of attacks	Features may need to be filtered or customized to reduce the training time
[20]	To dynamically configure the intrusion detection system (IDS) based on the vehicles location	Back propagation neural network and blockchain	Local samples (Simulated results)	IDS used in the small regions with micro blockchain.	It may require computation and configuration overhead for the large networks
[21]	To identify the malicious vehicle and broadcast the messages about the malicious vehicles	Convolutional neural networks	Simulated results	It avoids the unnecessary collisions	Various attack detection and mitigation techniques may need to be examine with different patterns
[22]	To detect the malicious node using hypothesis testing technique	Deep learning	NSL-KDD	Speedily detect the malicious vehicles	The evaluation is not considered for different traffic
[23]	The trust values to be shared between the vehicles and identifies the intrusions within the network	Support vector machine	Simulated results	Every node is aware of the next hop in case of any malicious node in the network and acting accordingly	It may be depends on few parameters to calculate the trust value
[24]	To provide the IDS based on the transfer learning method for in-vehicle network	Hybrid approach (Combining convolutional neural networks (CNN) and long short-term memory (LSTM))	Car hacking dataset and Defense Challenge 2020	Training and testing time reduced almost 30% compared to the previous methods	It may not be compatible to controller area network (CAN) networks under the various malicious nodes
[25]	To make on effective CAN bus system attacks	LSTM	Car hacking attack & defense challenge 2020	It can easily discriminating the normal and attack pattern	It requires the fine-tuning of the hyper parameters
[26]	To make the reliability and trustworthiness of federated learning (FL) process using blockchain	Federated learning	UNSW-NB15 dataset	It provides a decentralized secure reputation for intelligent transportation system	It may be limited to the particular dataset and need to consider the scalability issues
[27]	To detect and classify the attack in the CAN bus networks	Ensemble Model	CAN intrusion dataset	The lack of effectiveness of the conventional security measures are solved	Utilization of multiple models may requires the computation overhead

overall security framework's adaptability and responsiveness.

A. Motivation

An imperative requirement for robust security measures has arisen due to the expanding prevalence of autonomous and connected vehicles and the increasing integration of intelligent transportation systems. V2X networks are susceptible to a multitude of security risks due to their connectivity, which could compromise road safety through message tampering, impersonation, and potential attacks on critical components. This research addresses these difficulties by presenting a comprehensive security framework that utilizes modern technologies such as machine learning, and intrusion detection systems. The proposed solution relies on the implementation of virtual honeypots in RSUs functioning as decoys. These are utilized to entice and detect potential threats within a regulated setting. Moreover, the incorporation of ML improves the system's ability to recognize patterns of both regular and harmful behavior inside the V2X network. In essence, this undertaking will address the research questions (RQ) that have been delineated:

- **RQ1:** *How can a honeypot environment be made to effectively imitate real-world V2X communication using virtual RSUs?*
- **RQ2:** *How can potential security hazards, such as malicious activities and unauthorized access, be identified and responded to by the proposed IDS?*
- **RQ3:** *How to recognize and differentiate patterns of normal and malicious behavior that are seamlessly incorporated into the security framework of the V2X network using ML?*
- **RQ4:** *Which machine learning approaches are suitable for detecting threats in real-time and improving the system's ability to adapt and respond to new security incidents?*

B. Contribution

The main contributions of our study are as follows:

- A novel security framework that integrates the features of honeypots, virtual RSU, intrusion detection systems, machine learning, and edge computing has been proposed and evaluated.
- Our framework attains adaptive threat detection capabilities by incorporating machine learning methods. The machine learning algorithms acquire knowledge and identify patterns of both normal and malicious behaviors inside the V2X network, facilitating immediate identification of potential threats and a proactive approach to addressing security incidents.
- Our framework periodically exchanges information between PotRSU and actual RSUs in conjunction with cloud technology thus enabling the dissemination of knowledge about the attacker's behavior. This enhances global comprehension of new risks in V2X networks, promoting a collaborative and knowledgeable security community.
- Our proposed framework gathers information from the PotRSU and evaluates performance metrics across various machine-learning models. By prioritizing superior

accuracy and other performance indicators, we determine the most suitable model for detecting new security incidents within our proposed system.

The subsequent sections of this article are structured in the following manner: Section II knowledge about the preliminaries needed to understand the framework. Section III covers the state-of-the-art of the related works. Section IV provides an in-depth examination of the proposed system. Section V concentrates on presenting the outcomes and comparing the system's performance. Ultimately, Section VI concludes our work.

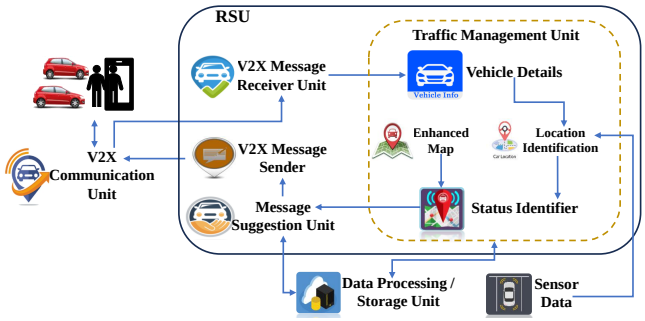


Fig. 2: General structure of RSU.

II. Preliminaries

A. Honeypots

A honeypot is a deliberately configured decoy system or network element designed to entice potential attackers [28]. By assuming the appearance of a susceptible target, it distracts malicious actors from legitimate systems and services [29]. Honeypots contribute to the improvement of V2X security by enticing potential threats into a regulated setting where their activities are meticulously observed and assessed. This enables security experts to acquire knowledge regarding emergent threats, methods employed by attackers, and weaknesses that are unique to V2X networks. By simulating a variety of V2X components, including roadside units and vehicles, they are capable of actively engaging with potential attackers [30]. This dynamic exchange facilitates the collection of up-to-date information regarding attempted attacks, allowing organizations to efficiently adjust and strengthen their security protocols. Honeypots are purposefully set up to deflect potential attackers' attention from the vital parts of the V2X networks. They function as appealing dummy targets, thereby serving as an early warning system and effectively mitigating the likelihood of triumphant attacks on the authentic, critical systems encompassed within the network. Implementing this prompt reaction mitigates the consequences of potential attacks on V2X networks and maintains secure and dependable communication among vehicles, infrastructure, and other relevant parties.

B. RSUs in V2X Communication

RSUs are integral components of V2X communication networks and serve a wide range of functions. These infrastructure elements collect real-time data from a variety of sources, including roadside sensors, traffic signals, other vehicles, and other roadside units, to function as data aggregators [10], [31]. Following this, the data is disseminated to adjacent vehicles, thereby substantially augmenting their situational awareness. They synchronize traffic signs, signal phases, and traffic lights in response to changing traffic conditions, thereby optimizing traffic flow.

The components of the RSU are illustrated in Figure 2. The RSU sends and receives messages and information to/from the V2X nodes, as indicated by the V2X communication unit. It is responsible for the collection of exhaustive data on adjacent vehicles, including but not limited to their make, model, and other distinctive characteristics that facilitate vehicle identification. The RSU contains sophisticated digital map data that enables users to locate and comprehend routes, navigation, and the road network in the vicinity of V2X entities. The precise location of V2X entities can be determined and accessed by roadside units via their location identification unit. The message suggestion unit generates status messages for dissemination to connected vehicles. This is achieved by utilizing the status identifier provided in the roadside unit to determine the current state of the V2X entities. It is the responsibility of the message-transferring unit of the roadside unit to transmit data and messages to nearby vehicles via V2X communication protocols [32].

C. Intrusion Detection System (IDS)

IDS are security technologies designed to protect against unauthorized or malicious activities by monitoring network or system activity. If such activities are detected, administrators are notified through the alert system or automated processes [33], [34]. The proposed IDS structure for ensuring security in V2X Networks is illustrated in Figure 3. It collects input patterns from V2X entities such as vehicles, pedestrians, RSU, and PotRSU. The collected information undergoes processing in the data processing unit, where redundant data and zero values are removed by the data cleaning unit. To enhance model performance, only relevant features are selected by the feature selection unit. The min-max approach is then employed for accurate results in the normalization part. Once the data is processed, the collected patterns are compared with predefined malicious patterns for identification, and patterns are classified based on behavior. Anomaly detection involves monitoring the communications and behavior of vehicles, aiding in identifying distinct attack patterns like those associated with denial of service (DoS), malware, or known intrusion attempts.

IDS also performs traffic analysis, scrutinizing network traffic to detect and indicate suspicious or malevolent packets. This function is crucial for overseeing communication among roadside infrastructure, vehicles, and central control systems. Real-time alerts are generated if suspicious activity is detected, enabling prompt action to mitigate potential security threats in vehicular networks. Consistent maintenance and updates are critical for repelling ever-evolving threats [35]. Additionally, it aids in detecting and resolving human-machine interaction issues, such as unauthorized access or tampering with vehicle control systems, and monitors the utilization of secure communication protocols.

III. Related Work

The safeguarding of V2X is a matter of great significance that has been extensively studied in recent years. Identifying

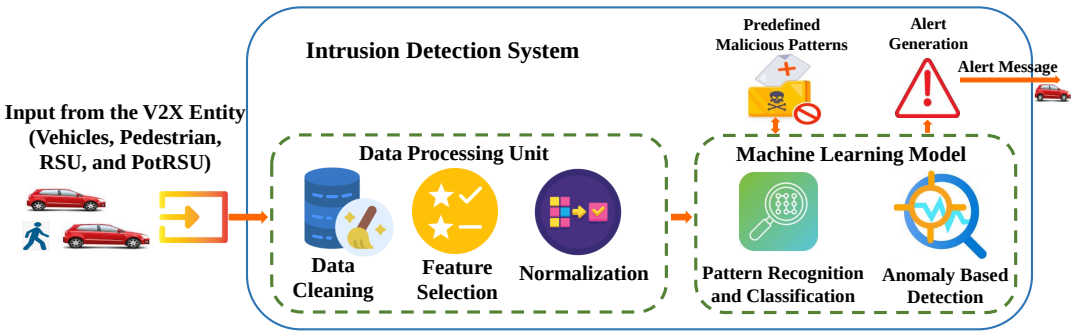


Fig. 3: Workflow of IDS in V2X.

the presence of unauthorized vehicles is a key challenge in this context. Employing a honeypot is also a crucial aspect of enhancing security, as it involves identifying malicious vehicles through the analysis of their behavior and attack strategies. Pashaei et al. proposed a technique that detects man-in-the-middle and distributed denial of service (DDoS) attacks using honeypot technology for industrial control systems (ICS). The authors suggested employing the Markov decision process (MDP) to design the honeypot, aiming to protect the network from unidentified attackers by analyzing their activities. To enhance the accuracy of the proposed system, the authors utilized classification and environmental agents. The classification agent learns the most complex policies to increase learning capability and enable early detection through analysis of potentially malicious behavioral patterns [36]. Baldo et al. have developed a honeypot-based electric vehicle supply equipment, enabling user interaction through the dashboard. This system enhances the effectiveness of the electric vehicle supply equipment (EVSE) device, contributing to potential improvements in future EVSE technology [37].

Singh et al. reviewed various IDSs employed in vehicular ad hoc networks (VANETs), exploring the associated benefits and challenges. The authors proposed the use of a honeypot to enhance the detection rate and bolster VANET security. They assert that employing honeypots improves VANET security, taking into account factors such as the number of nodes in different locations and the types of IDSs in use [38]. Prathapani et al. proposed an intelligent honeypot for detecting block-hole attackers in wireless mesh networks. In this approach, the intelligent honeypot functions as a detection agent and issues timely alerts in case of an attack. This method enhances the detection rate and reduces the false positive rate. However, it's important to note that Prathapani et al. specifically concentrated on addressing block-hole attacks and not on all other types of attacks [39]. Verendel et al. identified that wireless communication technology is more complex in vehicular communication and is susceptible to potential attacks. The authors employed honeypot technology to safeguard wireless communication within and between vehicles. The data has been collected using honeypot technology, and it is subsequently processed and analyzed by a central processing that is controlled by the operator. This analyzed information proves valuable in identifying attackers' behavior, understanding their techniques, and fortifying the defense against communication attacks in the future [40].

Anastasiadis et al. proposed a technique that utilizes honeypots to emulate sensors found in the internet of vehicles (IoV). They gather logs from the honeypot, including sequential patterns that capture attack propagations. The honey farm data undergoes analysis through a Markov chain model, and graph-based algorithms are employed to train models for identifying sequences of attack patterns from the honeypot data. This proposed technique proves effective in identifying common attacks and determining the geolocation of the attacker [41]. Zhang et al. employed three types of honeypots to detect malicious behaviors from attackers. They utilized medium and high interaction honeypots to identify CVE-2017-17215 attacks, which target universal plug-and-play (UPnP) router services. Additionally, a multi-port honeypot is implemented to enhance the honeynet's capacity. The proposed system effectively identified and captured unknown malicious attacks [42].

IV. Proposed Methodology

The process of configuring the Edge-enabled virtual honeypot-based intrusion detection system (EVHIDS) is a systematic and all-encompassing procedure that strengthens security in V2X networks, as illustrated in Figure 4. The process commences with the roadside unit being configured to function as a valid access point for authenticated entities in the V2X network. This guarantees a secure connection exclusively for authorized users. A virtual RSU (PotRSU), is deployed to imitate an actual RSU. PotRSU functions as a honeypot, enticing potentially malicious entities that are proactively in search of unsecured connections. IDS, which is strategically located near the PotRSU and acts as a monitor for network traffic analysis.

When malicious vehicles attempt to establish connections with PotRSU, the IDU collects essential information regarding their communication strategies, behavioral approaches, and attempted connections. Preprocessing is performed on this data to refine and extract relevant features, such as communication patterns and anomalies, thereby preparing it for comprehensive analysis. The IDU incorporates an advanced detection engine that examines the preprocessed data by utilizing a variety of detection techniques, including anomaly detection methods based on machine learning.

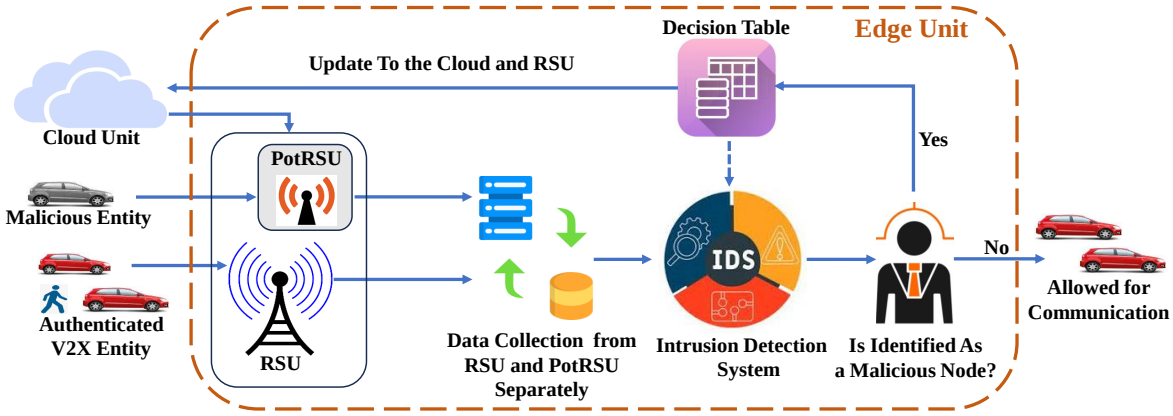


Fig. 4: The EVHIDS architecture.

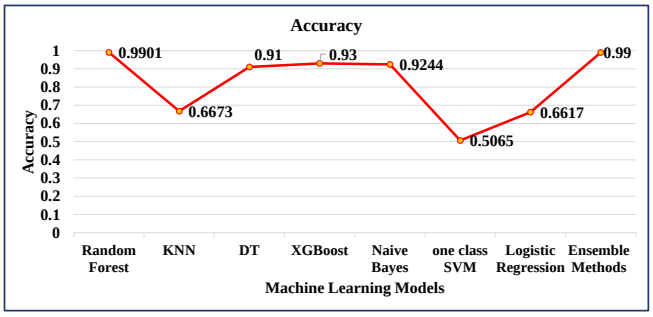
This technique for machine learning is capable of detecting anomalies that could potentially indicate attacks and can also adjust to evolving attack patterns. Subsequently, anomalies and detected patterns are transmitted to the cloud for storage, analysis, and dissemination among authorized parties. In the cloud, identified patterns are referenced in the form of a decision table or database, which serves to inform subsequent security measures. The actual roadside unit is not left behind; it gets information on patterns found and vehicle behavior that may be malicious. This update process provides the actual roadside unit with enhanced capabilities to identify and address potential hazards. The sequential steps and operations that comprise the entire proposed system are elaborated by Algorithm 1.

A. The data collection and preprocessing

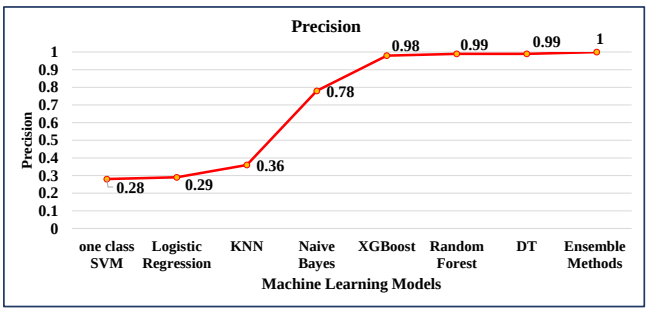
The feature description and sample collected data are shown in Table II and Table III respectively. Entities trying to connect to the PotRSU are treated as malicious entities in the proposed model. The data from PotRSU is forwarded to the IDS for advanced processing and behavioral analysis. We employ an ML algorithm to analyze behavior and detect attacks to accomplish this goal. To minimize the amount of time and resources needed, we further optimize the dataset by selecting features using data-cleaning techniques. We evaluated our proposed system on a 12th Gen Intel® Core™ i7-12700 CPU @ 2.10GHz, with 32 GB of RAM, running Windows 11 Pro 64-bit. The implementation of our model and feature engineering was done using veins-5.2 and Python 3.12. The data preprocessing phase employed by our work involves the following operations:

1) Data cleaning: As the first step, we removed redundant data and unnecessary characteristics with zero values. Then transformed representations that aren't numerical into numerical values. This is an important stage since numerical feature vectors are needed for ML algorithms. As a result, each vector—such as position, heading, acceleration, and speed, is converted into three numerical values: x, y, and z.
2) Feature selection: It's a technique for reducing data that has an impact on model performance by shortening the training period. We used matrix correlation and the Random Forest approach to choose pertinent characteristics. StartTime, SenderID, MessageID, Position x, position y, and speed are among the features that have been chosen.
3) Normalization: Data standardization is required for optimizing algorithms. As part of normalization, we converted every value into the range [0,1] using the min-max scalar approach.

The system encompasses alert generation and response mechanisms in addition to detection capabilities. When particular patterns are identified, the IDU can quickly produce warnings, allowing security administrators to take rapid action, which may include notifying authorities. The data that has been gathered and the patterns that have The data that has been gathered and the patterns that have been identified are of great value when it comes to conducting comprehensive research and analysis on the constantly changing threat environment within V2X networks.



(a)



(b)

Algorithm 1 Procedure for the EVHIDS system

procedure MALICIOUS VEHICLE DETECTION($Ingress\_Traffic$)
Initialize the V2X environment with PotRSU
- Deploy PotRSU at the edge that mimics RSUs.
- Set up the data collection unit and IDS unit.

Setting up data collection unit
(a)**If** a V2X node requests the PotRSU to establish a connection:
- Establish connection with PotRSU.
- Display the "Connection Established" message.
- Start capturing communication packets between vehicles and PotRSU.

EndIf
(b)**If** no request is received from V2X nodes or if the request is for RSU:
- Capture patterns and data from the RSU.
- Store captured data.

EndIf
Repeat steps 5(a),12(b) to continue listening for requests and capturing data

Preprocess the data
(a) Clean the data by removing noise and irrelevant information.
- Apply techniques such as smoothing, outlier detection, and thresholding to remove noisy data points.
- Remove categories that are not relevant to the analysis or occur infrequently.
-Use statistical methods or machine learning algorithms to detect and remove outliers from the dataset.

(b) Transform the data into a suitable format for machine learning algorithms.

Feature selection
(a) Extract relevant features from the pre-processed data. Features include vehicle identifiers, message types, communication frequency, etc.

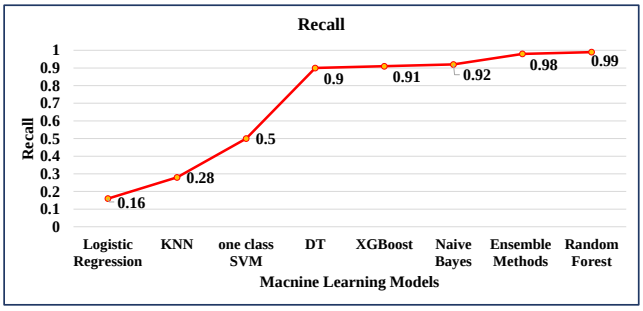
Perform Random Forest modeling
(a) For each tree in the forest
-Randomly sample a subset of features from the dataset
- Train a decision tree classifier using the sampled features and a subset of the training data

Repeat 6(a) for n_trees times to create an ensemble of decision trees.
(b) For each data point in the testing set:
-Pass the data point through each decision tree in the ensemble.
-Aggregate the predictions from all trees using a voting mechanism for classification.

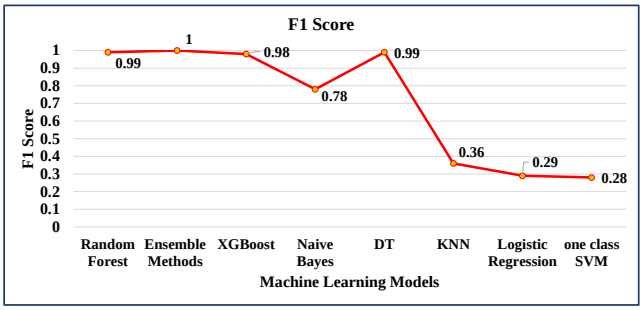
Deploy the model
-Integrate the trained ML models into the IDS components at Edge.

IDS component
(a) Continuously monitor V2X communication in real-time.
(b) If intrusion or misbehavior is identified
- Generate alerts
- Update the decision table in the edge unit as well as in the Cloud

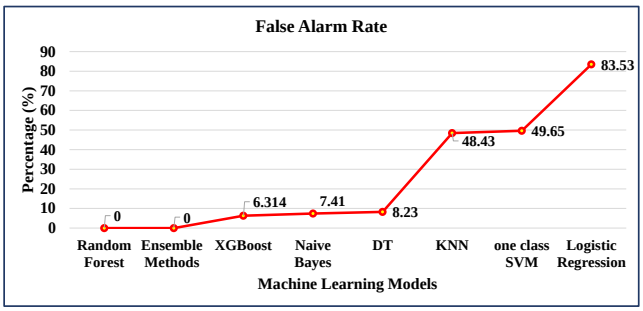
end procedure



(c)



(d)



(e)

Fig. 5: Validating the efficacy of the dataset using various models.

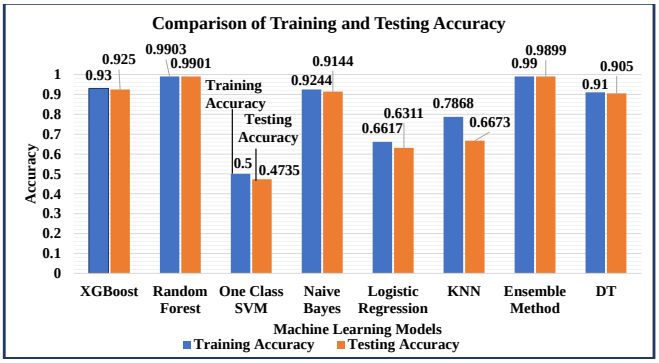


Fig. 6: Training and testing accuracy of ML models.

V. Results and Discussion

A. Evaluation Metrics

When evaluating the effectiveness of an IDS, it is crucial to employ a variety of metrics to obtain a comprehensive understanding of its performance. Different metrics can highlight various aspects of the system's capabilities and limitations.

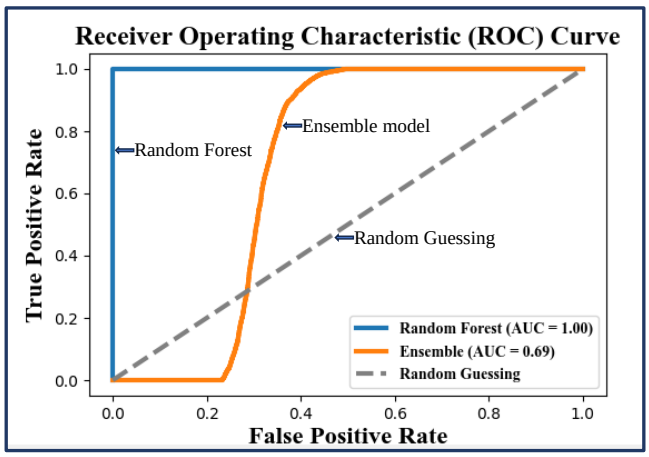


Fig. 7: The ROC of ML models.

TABLE II: Feature descriptions of the proposed model.

Features	Description	Type
Start Time	Vehicle starting time	Double
Sender ID	Sender Identification	Double
Position X	The geographic position of the X coordinate	Double
Position Y	The geographic position of the Y coordinate	Double
Speed	Vehicle Speed	Double
Message ID	Message Identifier	Double
Message Type	Type of message a vehicle sends	integer
Attacker Type	Kind of Vehicle attacking	Integer
Distance	Vehicle traveled length	Double

Key metrics to consider include detection rate, precision, accuracy, F1 score, and false alarm rate (FAR). These metrics help in determining how well the IDS identifies intrusions, the balance between correct and incorrect detection, and its overall reliability. Below, we delve into each of these evaluation metrics in detail to better understand their significance and application in assessing IDS performance.

- *Precision:* The IDS's accuracy in classifying an instance as an attack

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

- *Accuracy:* The overall correctness of the IDS.

$$Accuracy = \frac{TruePositives + TrueNegatives}{TotalInstances}$$

- *F1 Score:* A balanced measure is provided by the harmonic mean of recall and precision.

$$F1 = \frac{2 \times PrecisionRecall}{Precision + Recall}$$

- *FAR:*It calculates the percentage of negative cases that the model mistakenly classifies as positive.

$$FAR = \frac{FalsePositives}{TrueNegatives + FalsePositives}$$

- *Detection Rate (Sensitivity):* The proportion of actual attacks that the IDS correctly detects.

$$Sensitivity = \frac{TruePositives}{TruePositives + FalseNegatives}$$

- *Area under the ROC Curve (AUC-ROC):* The trade-off between true positive rate and false positive rate is indicated by the area under the receiver operating characteristic (ROC) curve.

B. Results and Inferences

To identify a suitable ML classifier for our proposed system, we trained and tested various models using the collected simulated data. We evaluated the F1 score, recall, accuracy, FAR, and precision for the machine learning models. Figure 5 presents a comparison of the accuracy, precision, recall, F1 score, and FAR of several machine learning models. It illustrates the comparative analysis of the performance of different machine learning models in terms of their accuracy. A higher position on the Y-axis indicates superior accuracy, implying that the random forest and ensemble models have successfully classified a larger proportion of instances correctly. Further, it can be observed that logistic regression (LR), one-class support vector machines (1-SVM), and K-nearest neighbor (KNN) produce less than 90% accuracy, while Naive Bayes, XGBoost, and decision tree (DT) consistently demonstrate more than 90% accuracy (Figure 5(a)).

Figure 5(b) facilitates a comparative examination of machine learning models based on their precision scores. A higher position on the Y-axis signifies superior precision, indicating that the models effectively minimize false positive predictions while maximizing the accuracy of positive predictions. Ensemble models and Random Forest models consistently exhibit high precision in the proposed system and problem domains, showcasing their reliability in correctly identifying positive instances with minimal false positives. Other models like LR, KNN, and DT exhibit lower precision, suggesting a propensity for false positive errors or misclassifications. Figure 5(c) illustrates a comparison based on recall scores. Ensemble and Random Forest models consistently exhibit 99% and 98% recall, respectively, demonstrating their efficacy in correctly identifying positive instances with minimal false negatives. Identifying disparities in recall scores among various models provides valuable insights into their respective performance characteristics and suitability for specific classification tasks.

Observations from Figure 5(d) reveal variations in F1 scores across different machine learning models. XGBoost and Random Forest models consistently demonstrate high F1 scores for the proposed system, indicating their robustness in achieving both high precision and high recall simultaneously. Observations from Figure 5(e) reveal variations in false alarm rates across different machine learning models. Random Forest and Ensemble models consistently exhibit low false alarm rates across diverse datasets and problem domains, indicating their reliability in correctly identifying negative instances without generating excessive false positive errors. It can also be studied that KNN, DT, Naive Bayes, and LR demonstrate higher false alarm rates, suggesting potential challenges in achieving a balance between sensitivity and specificity. Figure 8 illustrates the summary showcasing all evaluation metrics. Our analysis could also illustrate that Random Forest yields 100% accuracy, while the ensemble technique yields 69% accuracy when comparing the receiver

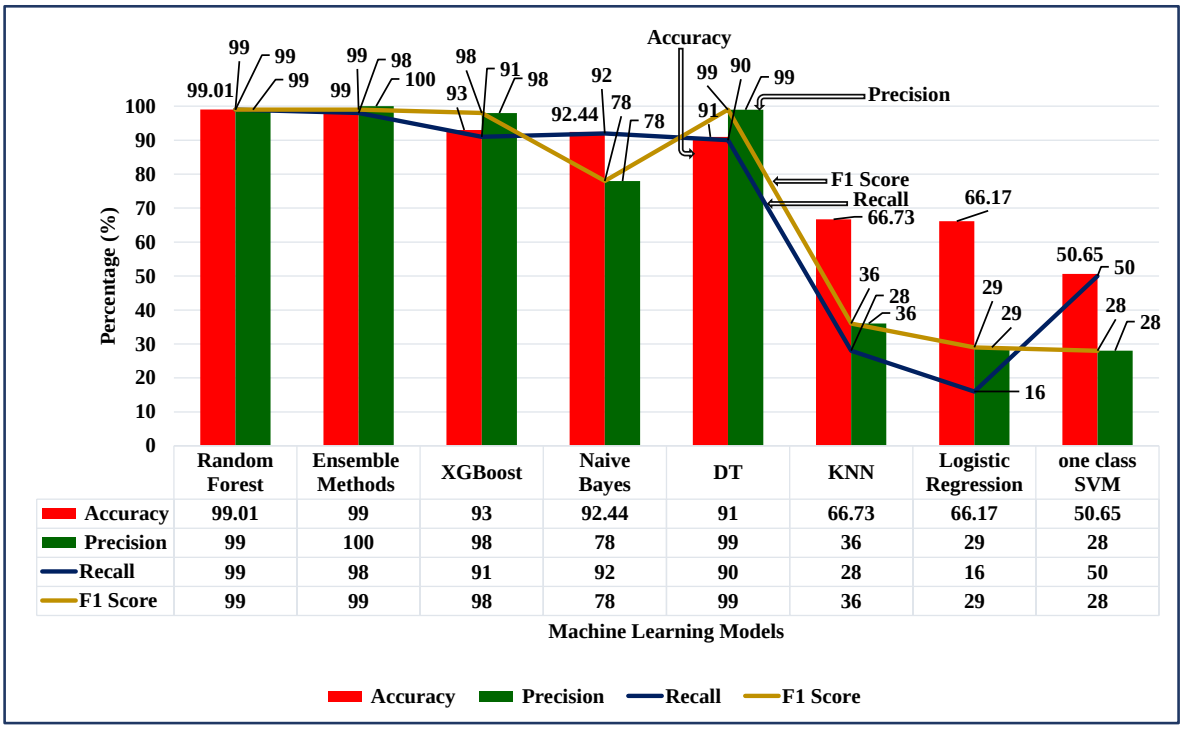


Fig. 8: Summary of evaluation of simulated data using different models.

TABLE III: Sample collected data of our proposed model.

Event ID	Start Time	Sender ID	Message ID	Position	Position Y	Speed	Message Type	Attacker Type	Distance
117	84.000003028978	96	433	1650.8062082113	2450.6248001783	13.998578815588	0	0	1953.3239315228
118	84.000003065903	102	434	1578.4613427264	2120.9360540062	13.999785439738	0	0	1939.0937131065
119	84.000003124501	108	435	1659.3938795371	2450.4780716181	13.999869057706	0	0	1923.3887349534
120	84.000003185663	114	436	1578.4506684253	2120.5707901428	13.999121439538	0	0	1904.622529885
121	84.000003403311	120	437	1672.1913562855	2450.659299592	13.999107769125	0	0	1884.5697129513
122	84.000003445943	126	438	1578.4476927174	2120.9589145361	13.999044324374	0	1	1865.6182666987
123	84.00000349302	132	439	1680.9291924543	2450.5176811377	13.999889892975	0	0	1844.075495343
124	84.000003498109	156	443	1578.4548774119	2120.90069908	13.999109892272	0	0	1821.0371727861
125	84.000003524896	150	442	1696.8696804541	2450.2645866362	13.999095289605	0	0	1797.7156332141

TABLE IV: Analysis of the proposed system with different models.

Name of the Model	Dataset Used	Accuracy	Precision	Recall	F1 Score
KIDS Model [43]	CIC-IDS 2017 dataset	98.6	98.02	98.27	98.92
Deep Neural Network [44]	CAN Dataset	98.68	96.78	96.72	96.71
Stacking [27]	CAN intrusion dataset	98.5	98.7	98.5	98.5
KIDS-UIDS [43]	i-VANET dataset	98.6	98.2	98.5	98.6
BiGAN [45]	KDD99	89.5	83.6	99.4	90.8
BiGAN Extended Model [22]	NSL-KDD	92.15	96.1	96.1	96.1
Support Vector Machine [46]	CAN Bus Dataset	97.9	98	96	97
Our Proposed Model	Simulated Results	99.01	99	99	99

operating characteristics (ROC) of the two methods. Accordingly, we infer that the Random Forest approach is appropriate for our framework.

Figure 6 and 7 display a comparison of the training and testing accuracy of various ML classifiers and ROC. The accuracy comparison reveals that the Random Forest model performs well for our proposed system in both training and testing scenarios. The ROC curves for the Random Forest, Ensemble model, and Naive Bayes indicate that the Random Forest model achieves 100% accuracy. Our proposed method is further compared with the existing IDS dataset [50], It can be observed that our proposed work achieves an accuracy of 99.01% for the Random Forest classifier, while the existing work using the Random Forest classifier achieves 98.60% accuracy. Hence, our proposed method effectively detects malicious nodes present in the V2X network. The training and testing times for this model won't have an impact on the system's performance in our suggested work. To ensure that the training and testing phases of the ML model do not interfere with the real-time communication between RSU and other connected devices, they must be carried out independently of other system entities, such as

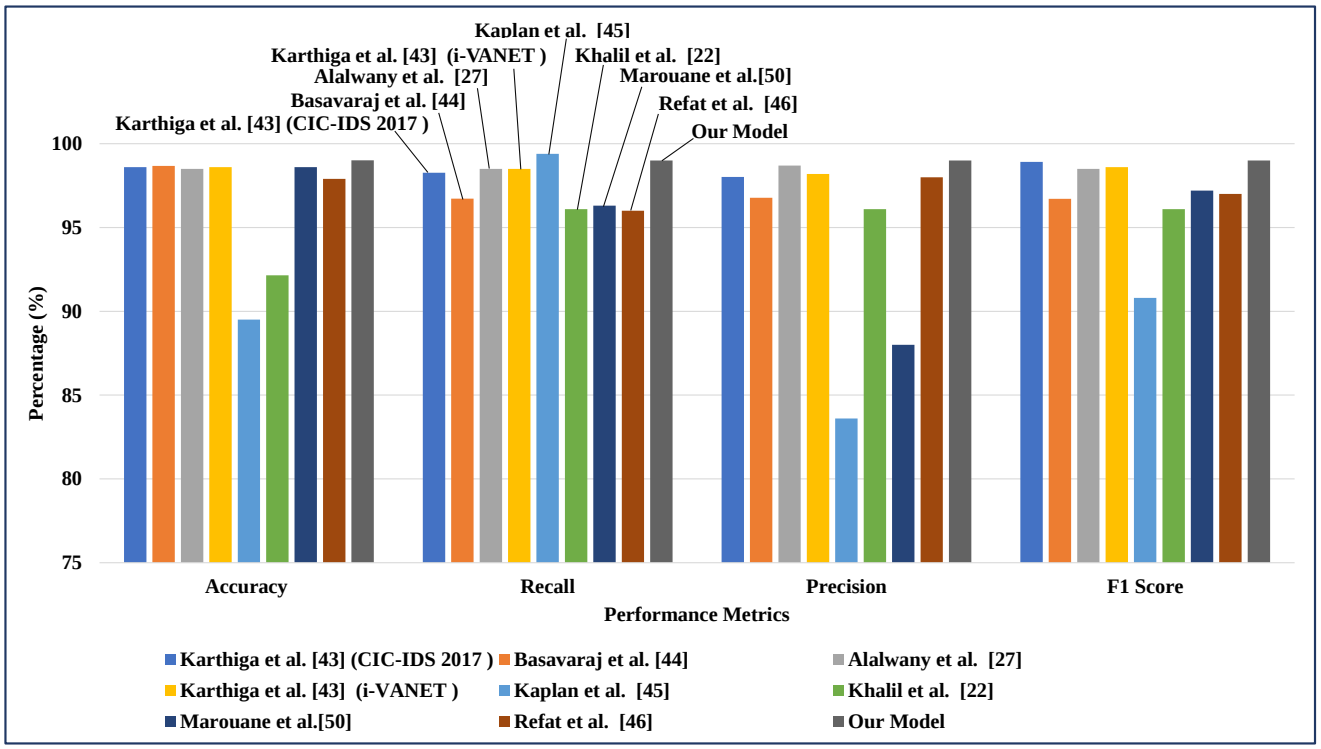


Fig. 9: Analysis of our proposed system with performance metrics.

TABLE V: Analysis of attack detection rate (%) of the proposed model with different attacks.

Reference	Brute Force attack	Botnet attack	DoS attack
[47]	92.7	95.9	96.7
[48]	92.8	97.9	98.1
[49]	90.9	97.3	96.3
[43]	93.9	98.9	98.5
Our model	98.8	99.02	99.22

TABLE VI: Analysis of the proposed system with performance metrics (%).

Reference	Accuracy	Recall	Precision	F1 Score
Karthiga et al. [43] (CIC-IDS 2017)	98.6	98.27	98.02	98.92
Basavaraj et al. [44]	98.68	96.72	96.78	96.71
Alalwany et al. [27]	98.5	98.5	98.7	98.5
Karthiga et al. [43] (i-VANET)	98.6	98.5	98.2	98.6
Kaplan et al. [45]	89.5	99.4	83.6	90.8
Khalil et al. [22]	92.15	96.1	96.1	96.1
Marouane et al. [50]	98.6	96.3	88	97.2
Refat et al. [46]	97.9	96	98	97
Our model	99.01	99	99	99

other vehicles and RSU. The IDU sends alert messages to the RSU and updates the patterns in the cloud for future reference. Table V presents a comparative analysis of our proposed model against other methods based on detection rates. The EVHIDS model demonstrates better detection rates for certain attacks, such as botnet attacks, denial-of-service attacks, and brute force attacks in the V2X environment. Our model is compared with various datasets and the results are shown in Table IV, Table VI, and Figure 9. The results exhibit significantly higher accuracy, precision, recall, and F1 score, which stand at 99.01%, 99%, 99%, and 99%, respectively.

VI. Conclusion

In conclusion, our proposed work introduces a pioneering approach to enhance vehicle security in V2X networks through the deployment of a PotRSU with an Intrusion Detection System. The amalgamation of virtual honeypot technology and machine learning, strategically positioned at the network edge, represents a significant leap forward in proactively identifying and mitigating potential risks in V2X communications. Our comparative analysis demonstrated the system's adeptness in distinguishing between malicious and regular behaviors, showcasing its adaptability to V2X scenarios. The results underscore the potential of our proposed method to elevate the security posture of intelligent transportation networks. By actively engaging and evaluating potential threats in a controlled environment, our system utilizes machine learning techniques to stay ahead of emerging threats, continuously adapting, learning, and improving over time. The study's findings underscore the importance of proactive security measures in V2X networks and the critical need for intelligent and flexible defenses against the ever-changing landscape of cyber threats. The knowledge gained from this research contributes to the broader discourse on safeguarding connected vehicles, fortifying resilience, and ensuring the integrity of next-generation intelligent transportation systems.

References

[1] A. Moubayed and A. Shami, "Softwarization, virtualization, and machine learning for intelligent and effective vehicle-to-everything

communications," *IEEE Intelligent Transportation Systems Magazine*, vol. 14, no. 2, pp. 156–173, 2020.

[2] F. Abbas, P. Fan, and Z. Khan, "A novel low-latency v2v resource allocation scheme based on cellular v2x communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2185–2197, 2018.

[3] Huaixu Gao, and Ying Tian, "Research on Road-Sign Detection Algorithms Based on Depth Network," *Engineering Letters*, vol. 31, no. 1, pp136-142, 2023.

[4] T. Deinlein, R. German, and A. Djanatliev, "5g-sim-v2i/n: towards a simulation framework for the evaluation of 5g v2i/v2n use cases," in *2020 European Conference on Networks and Communications (EuCNC)*. IEEE, 2020, pp. 353–357.

[5] C. Zhang, J. Wei, S. Qu, C. Huang, J. Dai, P. Fu, Z. Wang, and X. Li, "Implementation of a v2p-based vru warning system with c-v2x technology," *IEEE Access*, vol. 11, pp. 69 903–69 915, 2023.

[6] Lizhong Zhu, Xinfeng Yang, and Xianglong Huo, "A Study on Passenger Flow Control Scheme for Single-line Multi-station Urban Mass Transit Considering Passenger Flow Loss," *IAENG International Journal of Applied Mathematics*, vol. 54, no. 2, pp155-168, 2024.

[7] Changfeng Zhu, Yu Wang, Qingrong Wang, Jinhao Fang, Jie Wang, and Linna Cheng, "Research on Traffic Accident Prediction Based on KG-CWT-RGCNN-BiLSTM," *Engineering Letters*, vol. 31, no. 4, pp1402-1414, 2023.

[8] A. Alnasser, H. Sun, and J. Jiang, "Cyber security challenges and solutions for v2x communications: A survey," *Computer Networks*, vol. 151, pp. 52–67, 2019.

[9] H. Kong, W. Chen, S. Fu, H. Zheng, L. Du, and Y. Mao, "Obu design and test analysis with centimeter-level positioning for lte-v2x," in *2019 5th International Conference on Transportation Information and Safety (ICTIS)*. IEEE, 2019, pp. 383–387.

[10] S. Ma, F. Wen, X. Zhao, Z.-m. Wang, and D. Yang, "An efficient v2x based vehicle localization using single rsu and single receiver," *IEEE Access*, vol. 7, pp. 46 114–46 121, 2019.

[11] F. Hawlader, F. Robinet, and R. Frank, "Leveraging the edge and cloud for v2x-based real-time object detection in autonomous driving," *Computer Communications*, vol. 213, pp. 372–381, 2024.

[12] F. Luo and S. Hou, "Cyberattacks and countermeasures for intelligent and connected vehicles," *SAE International Journal of Passenger Cars-Electronic and Electrical Systems*, vol. 12, no. 07-12-01-0005, pp. 55–66, 2019.

[13] I. M. Varma and N. Kumar, "A comprehensive survey on sdn and blockchain-based secure vehicular networks," *Vehicular Communications*, vol. 44, p. 100663, 2023.

[14] M. Begum, G. Raja, and M. Guizani, "Ai-based sensor attack detection and classification for autonomous vehicles in 6g-v2x environment," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 4, pp. 5054–5063, 2024.

[15] S. Sharma and A. Kaul, "A survey on intrusion detection systems and honeypot based proactive security mechanisms in vanets and vanet cloud," *Vehicular communications*, vol. 12, pp. 138–164, 2018.

[16] Xiaobo Yang and Lianggui Liu, "Research on Traffic Flow Prediction based on Chaotic Time Series," *IAENG International Journal of Applied Mathematics*, vol. 53, no. 3, pp1007-1011, 2023.

[17] J.-H. Park, J.-w. Choi, and J.-S. Song, "How to design practical client honeypots based on virtual environment," in *2016 11th Asia Joint Conference on Information Security (AsiaJCIS)*. IEEE, 2016, pp. 67–73.

[18] M. Islabudeen and M. Kavitha Devi, "A smart approach for intrusion detection and prevention system in mobile ad hoc networks against security attacks," *Wireless Personal Communications*, vol. 112, no. 1, pp. 193–224, 2020.

[19] V. Praneeth, K. R. Kumar, and N. Karyemsetty, "Security: intrusion prevention system using deep learning on the internet of vehicles," *International Journal of Safety and Security Engineering*, vol. 11, no. 3, pp. 231–237, 2021.

[20] H. Liang, J. Wu, S. Mumtaz, J. Li, X. Lin, and M. Wen, "Mbid: Micro-blockchain-based geographical dynamic intrusion detection for v2x," *IEEE Communications Magazine*, vol. 57, no. 10, pp. 77–83, 2019.

[21] S. Anbalagan, G. Raja, S. Gurumoorthy, R. D. Suresh, and K. Dev, "Iids: Intelligent intrusion detection system for sustainable development in autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 12, pp. 15 866–15 875, 2023.

[22] A. Khalil, H. Farman, M. M. Nasralla, B. Jan, and J. Ahmad, "Artificial intelligence-based intrusion detection system for v2v communication in vehicular adhoc networks," *Ain Shams Engineering Journal*, vol. 15, no. 4, p. 102616, 2024.

[23] 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.

[24] N. Khatri, S. Lee, and S. Y. Nam, "Transfer learning-based intrusion detection system for a controller area network," *IEEE Access*, vol. 11, pp. 120 963–120 982, 2023.

[25] C. R. Kishore, D. C. Rao, J. Nayak, and H. Behera, "Intelligent intrusion detection framework for anomaly-based can bus network using bidirectional long short-term memory," *Journal of The Institution of Engineers (India): Series B*, vol. 105, p. 541–564, 2024.

[26] Z. Abou El Houda, H. Moudoud, B. Brik, and L. Khoukhi, "Blockchain-enabled federated learning for enhanced collaborative intrusion detection in vehicular edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 7, pp. 7661–7672, 2024.

[27] E. Alalwany and I. Mahgoub, "An effective ensemble learning-based real-time intrusion detection scheme for an in-vehicle network," *Electronics*, vol. 13, no. 5, p. 919, 2024.

[28] V. D. Priya and S. S. Chakkaravarthy, "Containerized cloud-based honeypot deception for tracking attackers," *Scientific Reports*, vol. 13, no. 1, p. 1437, 2023.

[29] S. C. Sethuraman, T. G. Jadapalli, D. P. V. Sudhakaran, and S. P. Mohanty, "Flow based containerized honeypot approach for network traffic analysis: An empirical study," *Computer Science Review*, vol. 50, p. 100600, 2023.

[30] I. Shaer, A. Haque, and A. Shami, "Multi-component v2x applications placement in edge computing environment," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*. IEEE, 2020, pp. 1–6.

[31] A. Haydari and Y. Yilmaz, "Rsu-based online intrusion detection and mitigation for vanet," *Sensors*, vol. 22, no. 19, p. 7612, 2022.

[32] Adebayo Kayode James, Aderibigbe Felix Makanjuola, Ibrahim Abdullahi Adinoyi, and Olateju Samuel Olaniyi, "On Formulation of the Vehicle Routing Problems Objective Function with Focus on Time Windows, Quantities and Split Delivery Priorities," *IAENG International Journal of Applied Mathematics*, vol. 51, no. 3, pp680-687, 2021.

[33] P. H. Mirzaee, M. Shojafar, H. Bagheri, T. H. Chan, H. Cruickshank, and R. Tafazolli, "A two-layer collaborative vehicle-edge intrusion detection system for vehicular communications," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*. IEEE, 2021, pp. 1–6.

[34] Yangke Yuan, Hong Dai, Zijian Wu, and Di Meng, "Novel Network Intrusion Detection Method Based on IPSO-MTWSVM Model," *Engineering Letters*, vol. 30, no. 2, pp892-897, 2022.

[35] Di Meng, Hong Dai, Qiaochu Sun, Yao Xu, and Tianwei Shi, "Novel Wireless Sensor Network Intrusion Detection Method Based on Light-GBM Model," *IAENG International Journal of Applied Mathematics*, vol. 52, no. 4, pp955-961, 2022.

[36] A. Pashaei, M. E. Akbari, M. Z. Lighvan, and A. Charmin, "Early intrusion detection system using honeypot for industrial control networks," *Results in Engineering*, vol. 16, p. 100576, 2022.

[37] M. Baldo, T. Bianchi, M. Conti, A. Trevisan, and F. Turrin, "Honeyevse: An honeypot to emulate electric vehicle supply equipments," in *European Symposium on Research in Computer Security*. Springer, 2023, pp. 145–159.

[38] S. Singh, S. Sharma, S. Sharma, O. Alfarraj, B. Yoon, and A. Tolba, "Intrusion detection system-based security mechanism for vehicular ad-hoc networks for industrial iot," *IEEE Consumer Electronics Magazine*, vol. 11, no. 6, pp. 83–92, 2021.

[39] A. Prathapani, L. Santhanam, and D. P. Agrawal, "Detection of blackhole attack in a wireless mesh network using intelligent honeypot agents," *The Journal of Supercomputing*, vol. 64, pp. 777–804, 2013.

[40] V. Verendel, D. K. Nilsson, U. E. Larson, and E. Jonsson, "An approach to using honeypots in in-vehicle networks," in *2008 IEEE 68th vehicular technology conference*. IEEE, 2008, pp. 1–5.

[41] M. Anastasiadis, K. Moschou, K. Livitckaia, K. Votis, and D. Tzovaras, "A novel high-interaction honeypot network for internet of vehicles," in *2023 31st Mediterranean Conference on Control and Automation (MED)*. IEEE, 2023, pp. 281–286.

[42] W. Zhang, B. Zhang, Y. Zhou, H. He, and Z. Ding, "An iot honeynet based on multiport honeypots for capturing iot attacks," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 3991–3999, 2019.

[43] B. Karthiga, D. Durairaj, N. Nawaz, T. K. Venkatasamy, G. Ramasamy, and A. Hariharasudan, "Intelligent intrusion detection system for vanet using machine learning and deep learning approaches," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, p. 5069104, 2022.

[44] D. Basavaraj and S. Tayeb, "Towards a lightweight intrusion detection framework for in-vehicle networks," *Journal of Sensor and Actuator Networks*, vol. 11, no. 1, p. 6, 2022.

[45] M. O. Kaplan and S. E. Alptekin, "An improved bigan based approach for anomaly detection," *Procedia Computer Science*, vol. 176, pp. 185–194, 2020.

[46] R. U. D. Refat, A. A. Elkhail, A. Hafeez, and H. Malik, "Detecting can bus intrusion by applying machine learning method to graph based features," in *Intelligent Systems and Applications: Proceedings of the 2021 Intelligent Systems Conference (IntelliSys) Volume 3*. Springer, 2022, pp. 730–748.

[47] K. Mahmood, J. Arshad, S. A. Chaudhry, and S. Kumari, "An enhanced anonymous identity-based key agreement protocol for smart grid advanced metering infrastructure," *International Journal of Communication Systems*, vol. 32, no. 16, p. e4137, 2019.

[48] I. Naqvi, A. Chaudhary, and A. Rana, "Intrusion detection in vanets," in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. IEEE, 2021, pp. 1–5.

[49] L. Nie, Y. Li, and X. Kong, "Spatio-temporal network traffic estimation and anomaly detection based on convolutional neural network in vehicular ad-hoc networks," *IEEE Access*, vol. 6, pp. 40 168–40 176, 2018.

[50] H. Marouane, A. Dandoush, L. Amour, and A. Erbad, "A review and a tutorial of ml-based mds technology within a vanet context: From data collection to trained model deployment," *Authorea Preprints*, 2023.