Received: 14 November 2023 | Revised: 18 December 2023 | Accepted: 11 January 2024 | Published online: 18 January 2024

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2024, Vol. 3(3) 307–315
DOI: 10.47852/bonviewJCCE42022066

BON VIEW PUBLISHING

A Machine Learning Model to Predict Cyberattacks in Connected and Autonomous Vehicles

Manoj K. Jha[1,*] **and Rishav Jaiswal**[2]

[1]*Department of Information Technology, University of Maryland Global Campus, USA*

[2]*Department of Civil Engineering, McMaster University, Canada*

Abstract: Connected and autonomous vehicles (CAVs) are largely at the experimental stage. Their successful deployment and field implementation require a careful consideration of their vulnerabilities to cyberattacks. The primary security vulnerability is in the controller area network (CAN) protocol, which permits communication among electronic control units in CAVs. To address this vulnerability and mitigate cyberattacks, machine learning (ML) algorithms can be developed for intrusion detection in CAV's CAN protocol. In this research, the data structure of certain experimental datasets on message injection attack from the Hacking and Countermeasure Research Lab is examined. A random forest classifier-based ML model is developed owing to its efficiency in predicting cyberattacks on CAVs consisting of over 3 million datasets. A number of procedures within the Python programming environment are employed to clean the dataset before performing the prediction. The prediction for intrusion detection is performed with a 70:30 split of the training: testing data with a random state of 11 and number of estimators as 200. The accuracy is found to be over 92% for all three scenarios in performing the prediction. The model can be deployed in real-time investigation of cyberattacks in CAVs if real-time data were available. The data cleaning method developed in this study can be applied in other ML applications consisting of large datasets, such as credit card fraud and drug discovery, to name a few.

Keywords: connected and autonomous vehicles, cyberattack, machine learning, random forest classifier, controller area network, data structure

1. Introduction

Connected and autonomous vehicles (CAVs) are largely at the experimental stage. A recent article by Lee and Hess [1] performed a survey of public perception toward CAVs. It was found that safety, privacy, and data security are key concerns for the slow adoption of CAVs.

CAVs use advanced communication technologies to establish real-time connectivity with other vehicles, infrastructure, and the environment. This connectivity makes it possible to share information about traffic patterns, potential risks on the road, and the best routes, improving overall traffic management and efficiency. CAVs rely on advanced sensors, algorithms, and artificial intelligence to function without any direct human input. They have the potential to increase traffic flow, decrease accidents brought on by human mistakes, and improve road safety by removing the need for human drivers. However, the incorporation of connection and automation in automobiles introduces new security concerns. Particularly vulnerable to cyberattacks is the controller area network (CAN) protocol, which facilitates communication among electronic control units (ECUs) in automobiles. CAN protocols could be used by malicious actors to launch cyberattacks that jeopardize the security, privacy, and functionality of the vehicle systems. These attacks can include RPM spoofing, gear spoofing, fuzzy attacks, denial of service (DoS) attacks, and others [2, 3]. Intrusion detection systems (IDS) are essential in CAVs in order to handle these security issues.

An IDS is a security tool made to find and respond to harmful or unauthorized activity on a system or network. Its main goal is to quickly detect potential security breaches or intrusions so that appropriate countermeasures can be taken to preserve the integrity of the system and limit future harm [4]. It is impossible to overestimate the importance of IDS in the context of CAVs. Because they rely on several communication technologies and intricate software systems as they grow more connected and autonomous, automobiles are more susceptible to hackers. These weaknesses can be used by burglars to gain entry without authorization, undermine the vehicle's performance, or even injure occupants and other road users directly.

Due to numerous security flaws in the CAN protocol, it has become a prime target for cyberattacks on CAVs. Malicious actors can manipulate CAN messages, introduce malicious commands, and interfere with communication networks, posing significant risks. Integration of IDS in CAVs enhances cyber risk identification and prevention by continuously scanning the network and system activity, detecting anomalies, deviations, or

*Corresponding author: Manoj K. Jha, Department of Information Technology, University of Maryland Global Campus, USA. Email: manoj.jha@faculty.umgc.edu

known attack patterns. As CAVs rely on sophisticated software, sensors, and communication networks, these components, lacking strong authentication and encryption in the CAN protocol, serve as potential entry points for cyberattacks, including DoS and spoofing attacks. External interfaces, such as infotainment systems and wireless modules, further expose vulnerabilities, leading to remote attacks, malware insertion, and unauthorized access. The consequences of cyberattacks on CAVs are severe, including compromised safety, service suspension, financial losses, harm to manufacturer's reputations, and a decline in public confidence. To mitigate these risks, it is crucial to address the CAN protocol's flaws, implement robust security measures, and deploy efficient IDS to ensure the development of reliable and safe CAV systems.

The absence of efficient IDS tailored for CAVs exacerbates the problem of cyberattacks. Traditional IDS techniques, including signature-based and anomaly-based methods, face limitations in dealing with the unique challenges presented by the CAN protocol. Consequently, there is a critical need for an innovative solution that not only comprehensively understands the intricacies of the CAN protocol but also leverages cutting-edge technologies to predict and thwart cyberattacks in real-time.

The motivation behind this research stems from the imperative to ensure the secure deployment of CAVs, considering the pivotal role they are expected to play in the future of transportation. Public perception, as revealed by Lee and Hess [1], has identified safety, privacy, and data security as major hurdles. This underscores the urgency to address the cybersecurity challenges confronting CAVs, particularly the vulnerabilities within the CAN protocol. The unique challenges lie in the dynamic nature of CAV environments, the lack of authentication and encryption in the CAN protocol, and the potential consequences of cyberattacks on passenger safety and overall system functionality.

This research makes a significant contribution by introducing a novel machine learning (ML) model based on a random forest (RF) classifier for intrusion detection in CAV's CAN systems. A significant amount of research has been done to evaluate the application of ensemble-based ML models for civil engineering applications [5]. In the current research, the primary objective is to develop a sophisticated ML model capable of accurately predicting cyberattacks, specifically targeting the vulnerabilities within the CAN protocol. The contribution lies in the formulation of a data cleansing methodology within the Python programming environment and the subsequent application of the RF classifier to datasets from the Hacking and Countermeasure Research Lab (HCRL).

The significance of this research extends beyond the realm of CAVs, with potential applications in other domains grappling with cybersecurity challenges in large datasets. The developed ML model not only enhances the security posture of CAVs but also introduces a robust methodology for addressing security concerns in diverse applications, such as credit card fraud detection and drug discovery. The implications of this research are profound, setting the stage for a more secure and trustworthy integration of CAVs into the broader transportation landscape.

We develop several procedures within the Python programming environment to clean the dataset before performing the prediction. We create three test scenarios of training and testing split (60:40, 70:30, and 80:20) of the datasets to examine the accuracy of the model. The procedures are described under Section 3.

In the subsequent sections of this paper, we delve into the core components and methodologies employed in our research. Section 2 provides literature review and an in-depth exploration of the CAN protocol vulnerabilities within CAVs and the motivation behind developing a specialized IDS. Section 3 elucidates the methodology, detailing the development of our innovative RF classifier, the careful examination of experimental datasets, and the rigorous testing procedures conducted to validate its efficacy. Meanwhile, section 3 offers a comparative analysis, highlighting the advantages of our proposed work over existing solutions in the field. Furthermore, Section 4 discusses the implications of our findings and their significance in enhancing cybersecurity for CAVs. Finally, Section 5 concludes the paper by summarizing key contributions, discussing potential avenues for future research, and emphasizing the broader impact of our work in the field of autonomous vehicle security.

2. Literature Review

2.1. Overview of CAN protocol in vehicles

The CAN protocol used in vehicles is fundamental for facilitating communication among ECUs. Reducing the complexity and expense of wiring inside the car is one of the primary reasons CAN was developed. Systems where only little information needs to be communicated are most suited for it [6]. Despite having been developed with the automobile industry in mind, it is now utilized in a variety of different industries and control systems, including manufacturing, medical devices, lifts, robots, building automation, and manufacturing [7]. However, the widespread use of CAN in modern vehicles also introduces security challenges that need to be addressed. In recent years, several studies have highlighted the vulnerabilities and potential security threats associated with the CAN protocol. The CAN protocol's lack of built-in security features is one of its main security issues. The original architecture of CAN placed less emphasis on implementing strong security measures and more on offering a dependable and effective means of communication. This absence of security features leaves the CAN network susceptible to various cyberattacks.

Walker [8] highlighted the flaws in the CAN protocol and showed how an attacker might take control of an automobile's ECUs and modify their behavior. By successfully executing attacks during their tests, such as inserting arbitrary CAN messages and changing the behavior of the car, the study demonstrated the necessity for stronger security safeguards in the CAN protocol. Another significant security difficulty is presented by the CAN protocol's absence of authentication and encryption. It becomes challenging to confirm the legitimacy and integrity of the messages transferred across the network in the absence of adequate authentication procedures. Due to the lack of encryption, CAN network data transmissions are vulnerable to eavesdropping and unauthorized access [9, 10]. Due to the lack of authentication in the CAN protocol, it is possible to masquerade an ECU or replace a legitimate ECU with a malicious one using a hardware device [11].

According to a recent analysis from Tencent's Keen Security Lab [12], hackers can leverage these loopholes to take complete control of a Tesla vehicle's infotainment system without the user's involvement. In one experiment, they got access via wireless Wi-Fi/Cellular network, hacked a number of in-vehicle systems, such as firmware on an IC or gateway, and inserted malicious messages into the CAN bus to carry out a number of tasks, including opening the car door, window, and trunk. Furthermore, the CAN protocol's security issues are made worse by the dynamic nature of the vehicle environment [13]. Vehicles frequently connect to and disconnect from the network, giving

potential attackers an opportunity to enter the system without authorization when the network is being reconfigured. The open and uncontrolled nature of the CAN network increases the risk of unauthorized devices or malicious ECUs gaining entry, leading to potential security breaches.

The literature emphasizes the security difficulties in the CAN protocol that automobiles encounter. The absence of built-in security features, exposure to attacks, such as message injection and spoofing, and the lack of authentication and encryption are some of these difficulties. To improve the CAN protocol's security and reduce potential cyber threats in CAVs, it is imperative to solve these issues.

2.2. IDS for vehicle networks

IDS are essential for reducing the risk of cyberattacks on vehicle networks. In the context of CAVs, several methods and procedures have been developed to identify intrusions and respond to them.

The signature-based detection method is widely used. Signature-based IDS rely on predefined patterns or signatures of known attacks [14]. These signatures were developed using the traits and tactics of well-known attacks. The IDS creates an alert to show a potential intrusion when a network event matches a signature. However, one drawback of signature-based detection is that it depends on a large and current signature database. It may be difficult for the IDS to accurately detect new and emerging attacks since they might not have well-known signatures.

An alternative approach is anomaly-based detection. Anomaly-based detection is a different strategy. Network traffic and system behavior are analyzed using anomaly-based IDS to provide a baseline of typical operations [15]. Anomaly detection is helpful for spotting new threats without established signatures. However, if the IDS is not trained on a representative dataset or if genuine network behavior dramatically deviates from the specified baseline, it might also produce false positives (FPs). A typical architecture of intrusion detection of CAN bus network is shown in Figure 1.

ML approaches have become more popular in IDS for vehicle networks in recent years. Large amounts of network data can be analyzed by ML algorithms to spot patterns and detect known and unidentified attacks [16]. Among the notable contributions in the field, the work by Zheng et al. [17], Zheng et al. [18], and Zheng et al. [19] highlights the significance of activation functions in deep convolutional neural networks and the application of deep learning in modulation classification. On labeled datasets, supervised learning algorithms, such as RF, SVM, and ANN, can be trained to identify the traits of different attacks and correctly classify them. Without relying on predetermined attack fingerprints, unsupervised learning algorithms can discover anomalous patterns and behaviors, such as clustering and anomaly detection.

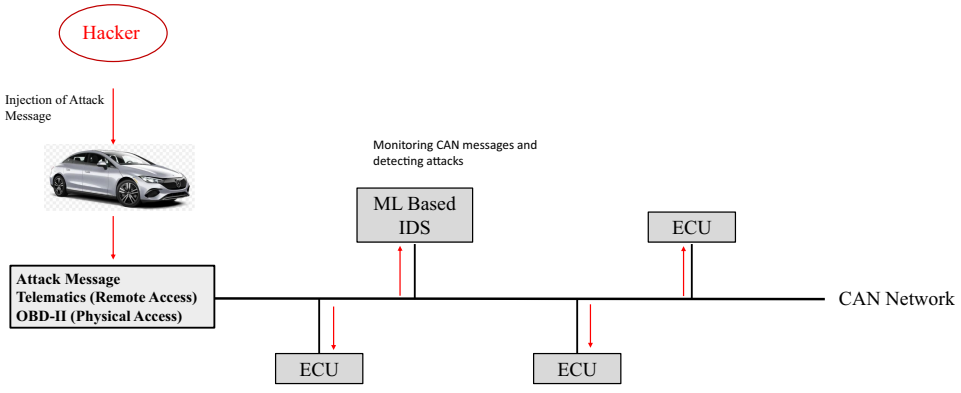
Traditional IDS techniques do, however, have several drawbacks. First, they frequently produce warnings for harmless network events due to high FP rates, which causes unneeded disruptions and extra research labor. Second, because they cannot analyze the payload of encrypted messages, IDS may have trouble processing encrypted or obfuscated network traffic. Third, because IDS approaches demand substantial computational resources for real-time monitoring and analysis, they may not scale well to large-scale car networks with numerous ECUs and enormous data volumes. Researchers have suggested cutting-edge solutions to get beyond these constraints, like hybrid approaches that integrate signature-based and anomaly-based detection techniques. Hybrid IDS maximize the benefits of both strategies to increase precision and decrease FPs. Furthermore, including ML algorithms into IDS has yielded promising results due to their ability to adapt to new attack patterns and minimize FPs through ongoing learning.

Existing approaches to intrusion detection for automotive networks include ML-based, anomaly-based, and signature-based techniques. These techniques have advantages, but they also have drawbacks in terms of precision, FPs, scalability, and capacity to handle encrypted traffic. Future studies should concentrate on creating more sophisticated and reliable IDS methods that meet these constraints and guarantee the safety of connected and autonomous cars.

2.3. ML for intrusion detection

ML techniques have shown significant potential in intrusion detection for vehicle CANs [20–24]. These methods use the capabilities of statistical models and algorithms to analyze network data and spot trends related to cyberattacks. To protect automotive CAN systems, ML in intrusion detection offers several benefits. For example, large amounts of complex data generated by vehicle networks may be rapidly analyzed and processed by ML algorithms, allowing for the discovery of trends and behaviors linked to cyberattacks [25]. Based on labeled datasets, supervised learning algorithms may correctly categorize network

Figure 1
Typical architecture of intrusion detection of CAN bus network



instances as benign or harmful. Examples of such include RF, support vector machine, and artificial neural network. Without depending on established attack fingerprints, unsupervised learning methods like clustering and anomaly detection can spot departures from the norm. ML-based IDS can detect new or zero-day attacks, adapt to changing attack patterns, and reduce FPs, thereby increasing the security of CAVs and the accuracy of CAN-based intrusion detection.

Cyberattacks on CAVs could result in the loss of personal information, physical injury, or even death, among other significant consequences. Threats to the CAN's bus system might be challenging to detect with conventional network intrusion detection systems. In order to safeguard the CAN bus in automobiles from infiltration, researchers have created intrusion detection models using ML approaches [26]. To understand the characteristics of attack behavior and categories threats in in-vehicle networks, Lin et al. [24] suggested an intrusion detection model based on the VGG16 deep learning classifier. Song et al. [3] developed a deep convolutional neural network-based model that recognized message injection threats by analyzing the sequential patterns of in-vehicle network traffic. The model was very effective and incorrectly classified very few messages as normal. For example, 12 out of 11,366 DoS attacks were classified normal, 36 out of 13,441 of fuzzy attacks were classified as normal, and 27 out of 19,862 spoofing gear message attacks were classified as normal.

Mansourian et al. [27] suggested a long short-term memory (LSTM)-based IDS for the CAN bus that uses the temporal correlations between messages to identify anomalies. It is specifically a one-class classifier that has been trained with data free from attacks to forecast the value of CAN messages in the future. The result showed a prediction accuracy of 0.977, 0.894, and 0.893 to detect attacked messages of type fuzzy, gear spoofing, and RPM spoofing, respectively. In order to avoid disastrous crashes and disruptive effects, real-time intrusion detection with little processing resources is required. In order to enable real-time intrusion detection in CAVs with the least amount of processing resources, Kumar and Das [9] suggested an intrusion detection approach based on logical analysis of data. The result showed a FP rate of 0.078.

Bari et al. [20] investigated the effectiveness of a ML-based IDS using support vector machine, decision tree, and K-nearest neighbor algorithms on real-world datasets. For the dataset provided, the models showed accuracy ranging from 93% to 99%. A LSTM-based IDS was proposed by Hossain et al. [22] to identify and counteract CAN bus network attacks. An effective IDS model was created by Basavaraj and Tayeb [21] utilizing a neural network to identify anomalies in the vehicular system. For the purpose of CAN security, Kalkan and Sahingoz [23] presented a ML-based IDS. These studies show how ML may be used to build reliable intrusion detection models for protecting the CAN bus system in CAVs.

Many researchers have developed models for intrusion detection in CAN with different datasets resulting into an accuracy of up to 0.99. For the dataset we have used in this study, the prediction accuracy is found to be over 0.92. The dataset we employed to detect cyberattacks on certain experimental CAV dataset was made publicly available by Seo et al. [2].

One of the underlying issues in ML-based approaches for CAV's security not sufficiently addressed in previous works is to catch true cyber intrusions and minimize instances of FPs. While applicable to the general class of unsupervised algorithms in the labeling context, Fang and Zhu [28] developed a formulation for an active learning paradigm with uncertain information. They offered an algorithm that used an error-reduction sampling estimation. The first author extended the active learning concept called modular active learning (modAL) for minimizing risk and uncertainty in transportation and construction scheduling [29].

This paper aims to provide a detailed exploration of our ML-based IDS model, offering insights into its development, dataset cleaning procedures, and evaluation metrics. The advantages of our approach, along with its implications for real-time investigation of cyberattacks in CAVs, are discussed. Our approach involves a meticulous analysis of experimental datasets, leading to the creation of a robust RF-based ML model. In comparison to existing works, our study stands out for its comprehensive examination of experimental datasets, totaling over 3 million instances, sourced from the HCRL. The adoption of the RF classifier reflects its efficiency in handling large datasets and predicting cyberattacks.

3. Research Methodology

We employ RF-based ML to analyze the cyberattacks on certain experimental CAV datasets of the HCRL obtained from the public dataset available through the Hacking and Countermeasure Research Lab (HCRL) website made public by Seo et al. [2] to foster further research. This dataset offers several advantages over other benchmark datasets. Firstly, it is helpful for intrusion detection research since it offers current behavioral traits and attack sequences of CAN signals. Secondly, it contains a vast array of attributes generated from the CAN IDs and DATA fields that accurately represent the content of network packets. Thirdly, it has several intricate features that help an ML model learn to distinguish between legitimate and malicious signals more precisely. The DoS attacks, fuzzy attacks, gear spoofing, and RPM spoofing are the four main attack types represented in the HCRL dataset. These attack datasets were created in a controlled setting by emulating ECUs and inserting fabricated CAN messages [2].

3.1. ML methodologies

The choice of RF in the present study is justified by its efficiency in handling large datasets and its ability to capture complex relationships within the experimental data. RF is an ensemble learning method that builds multiple decision trees during training and outputs the mode of the classes (classification) or the mean prediction (regression) of the individual trees. It is known for its robustness, ability to handle high-dimensional data, and resistance to overfitting. It also provides a feature importance ranking, aiding in the identification of critical variables. But caution should be exercised in employing an RF model since the training time might increase with a higher number of trees, affecting real-time applications.

3.2. Labeling procedure

The HCRL dataset includes four main attack types: DoS, fuzzy attacks, gear spoofing, and RPM spoofing. These attacks were carefully emulated in a controlled setting by injecting fabricated CAN messages, simulating ECUs behavior. Each CAN message in the dataset is labeled with a flag (T or R), where T represents

injected special attack messages, and *R* represents normal messages. This labeling process is crucial for training the classification algorithm to distinguish between normal and malicious behavior in the CAN system. Table 1 shows the shape of the raw data in the four categories.

Table 1
Shape of raw data

Data type	Number of rows	Number of columns
DoS	3,665,770	12
Fuzzy	3,838,859	12
Gear	4,443,141	12
RPM	4,621,701	12

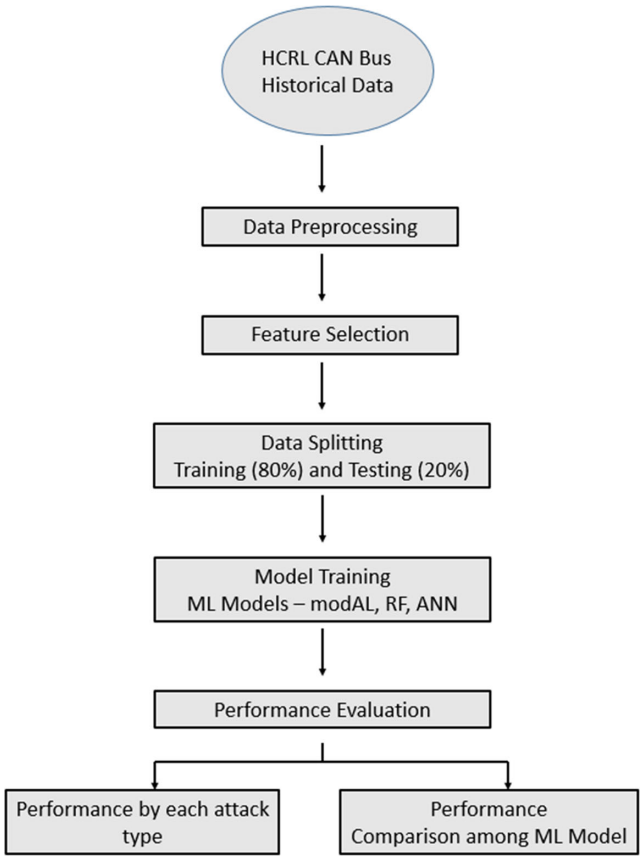
The data attributes for each CAN message, as summarized in Table 2, include the timestamp (recorded time in encoded form in seconds), CAN ID (identifier of the CAN message in hexadecimal format), DLC (number of data bytes, ranging from 0 to 8), DATA [0–7] (data value in bytes), and a flag indicating whether the message is a normal message or an injected attack message. The flag is either *T* or *R*; *T* represents injected special attack messages while *R* represents normal messages.

3.3. Data preprocessing

Prior to model training, a data cleansing algorithm was implemented in Python to handle missing or corrupted data and ensure the dataset's integrity. The algorithm removes rows with missing values and non-numeric data, resulting in a clean training dataset. Attributes such as timestamp, Can_Id, and DLC, which do not contribute significantly to prediction, were dropped to streamline the dataset.

Figure 2 shows the methodological framework.

Figure 2
Methodological framework



The features in the clean dataset include timestamp, CAN ID, DLC, and DATA [0–7]. While timestamp and CAN ID are crucial for chronological and message identification purposes, DLC and DATA [0–7] contain information about the data bytes. To improve model performance, correlation matrices were analyzed for each attack type dataset (DoS, fuzzy, gear, RPM). High correlation was observed between data pairs 3 and 4, suggesting a potential for feature engineering. However, due to the relatively quick prediction results and to preserve all attributes for comprehensive analysis, no features were dropped in this instance.

Tables 3, 4, 5, and 6 show the standard statistics of the clean training dataset in the four categories.

Table 2
Sample dataset with class labels

Timestamp	CAN ID	DLC	D0	D1	D2	D3	D4	D5	D6	D7	Flag
1.478198e+09	0440	8.0	ff	00	00	00	ff	d6	08	00	R
1.478196e+09	043f	8.0	10	40	60	ff	7d	96	09	00	R
1.478193e+09	02a0	8.0	20	00	95	1c	97	02	bd	00	R
1.478193e+09	02a0	8.0	20	00	95	1c	97	02	bd	00	R
1.478196e+09	0329	8.0	86	ba	7f	14	11	20	00	14	R
1.478191e+09	0370	8.0	00	20	00	00	00	00	00	00	R
1.478196e+09	02c0	8.0	15	00	00	00	00	00	00	00	R
1.478191e+09	0260	8.0	19	22	22	30	ff	8f	6f	1c	R
1.478193e+09	0370	8.0	00	20	00	00	00	00	00	00	R
1.478191e+09	0440	8.0	ff	00	00	00	ff	bc	08	00	R
1.478193e+09	043f	8.0	01	45	60	ff	6b	00	00	00	T
1.478193e+09	02a0	8.0	40	00	95	1c	97	02	bd	00	R
1.478198e+09	0316	8.0	05	21	74	09	21	20	00	6f	R
1.478198e+09	0430	8.0	00	00	00	00	00	00	00	00	R
1.478191e+09	0329	8.0	0c	b3	7e	14	11	20	00	14	R

Table 3
Standard statistics of the clean dataset for DoS

	Data_0	Data_1	Data_2	Data_3	Data_4	Data_5	Data_6	Data_7	Flag
Count	1177098	1177098	1177098	1177098	1177098	1177098	1177098	1177098	1177098
Unique	16	5	9	2	19	24	20	100	2
Top	00	00	00	00	00	00	00	00	R
Freq	907163	992397	1165921	1165921	1136146	1118503	1114199	1100048	589577

Table 4
Standard statistics of the clean dataset for fuzzy

	Data_0	Data_1	Data_2	Data_3	Data_4	Data_5	Data_6	Data_7	Flag
Count	646368	646368	646368	646368	646368	646368	646368	646368	646368
Unique	93	97	92	92	96	95	93	99	2
Top	00	00	00	00	00	00	00	00	R
Freq	342920	434534	616494	616413	587791	583778	579206	547500	646115

Table 5
Standard statistics of the clean dataset for gear

	Data_0	Data_1	Data_2	Data_3	Data_4	Data_5	Data_6	Data_7	Flag
Count	825542	825542	825542	825542	825542	825542	825542	825542	825542
Unique	49	34	44	5	30	30	63	82	1
Top	00	00	00	00	00	00	00	00	R
Freq	482600	587455	798833	798286	766156	735013	740571	715994	825542

Table 6
Standard statistics of the clean dataset for RPM

	Data_0	Data_1	Data_2	Data_3	Data_4	Data_5	Data_6	Data_7	Flag
Count	943906	943906	943906	943906	943906	943906	943906	943906	943906
Unique	16	14	26	3	20	24	20	74	1
Top	00	00	00	00	00	00	00	00	R
Freq	498754	607656	833320	833320	791534	772728	861589	747960	943906

Figures 3, 4, 5, and 6 show the correlation matrix for each of the four datasets.

The higher correlation among data pairs is demonstrated by a higher positive decimal fraction in Figures 3, 4, 5, and 6. In all four figures, data pairs 3 and 4 exhibit the highest correlation confirming that one of them can be dropped if computational burden becomes an issue. But, because we obtained the prediction results relatively quickly (in less than a minute), we decided to keep all of the attributes to perform the predictive analytics.

Figure 3
Correlation matrix for the DoS dataset

3.4. Evaluation metrices

Several assessment criteria are used in the proposed framework for intrusion detection in the CAV's CAN system to evaluate the effectiveness of the ML model. The model's precision, recall, and overall efficacy in identifying and categorizing cyberattacks are all quantified by these evaluation metrices. The main evaluation metrices used in this study are confusion matrix, precision, recall, accuracy, and F1-score.

Figure 4
Correlation matrix for the fuzzy dataset

Figure 5
Correlation matrix for the gear dataset

Figure 6
Correlation matrix for the RPM dataset

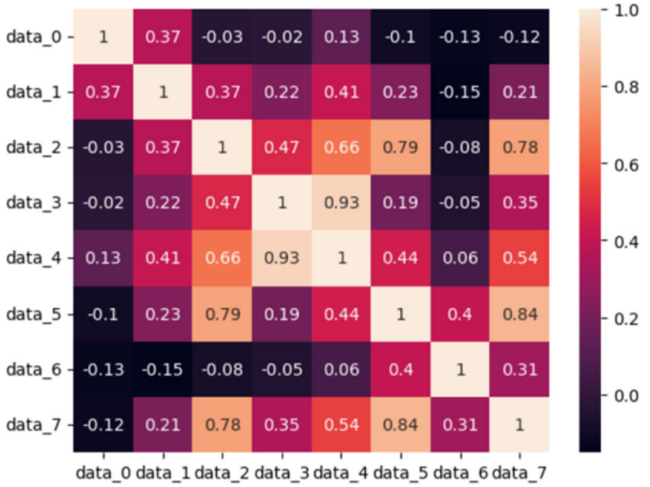
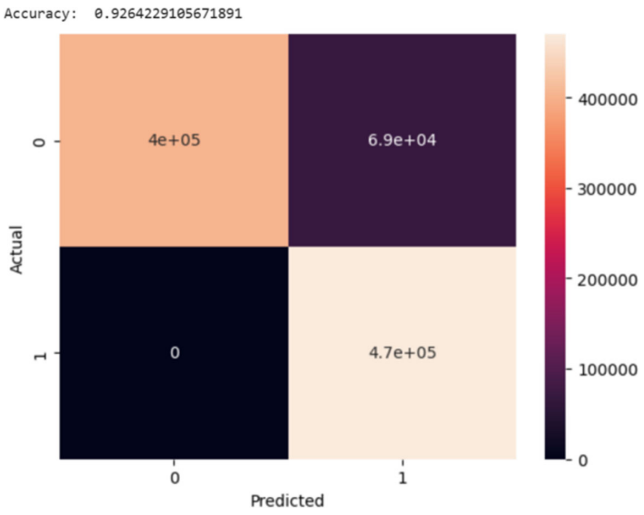


Figure 7
Confusion matric for the DoS dataset



3.4.1. Confusion matrix

The confusion matrix is a table that displays the counts of true positive (TP), true negative (TN), FP, and false negative (FN) predictions to summarize the performance of a classification model. It gives a thorough overview of the model's performance across many classes and aids in identifying the numerous kinds of errors the model makes.

3.4.2. Precision

Precision is the ratio of TP predictions to the total number of positive predictions made by the model. When the cost of FPs is significant, it is very helpful because it measures how well the model can identify positive cases. It is expressed as:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

3.4.3. Recall

The ratio of TP predictions to all the actual positive cases in the dataset is known as recall, often referred to as sensitivity or TP rate. It assesses how well the model can detect every positive event, even those that are overlooked (FN). It is expressed as:

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

3.4.4. Accuracy

Accuracy is the ratio of correctly classified instances (both TPs and TNs) to the total number of instances in the dataset. It gives a broad indication of how well the model predicts both positive and negative events. When the dataset is unbalanced or the cost of misclassifying distinct groups fluctuates, accuracy may not be the most reliable metric. It is expressed as:

$$Accuracy = \frac{Correct\ Predictions}{Total\ Predictions} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

3.4.5. F1-score

The harmonic mean of recall and precision is known as the F1-score. It is appropriate for imbalanced datasets since it offers a balanced metric that takes into account both precision and recall. The F1-score provides an overall assessment of the model's performance in terms of both FPs and FNs by combining the data from accuracy and recall into a single metric. It is expressed as:

$$F1\ Score = \frac{2\ X\ Precision\ X\ Recall}{Precision + Recall} \quad (4)$$

These evaluation metrices offer a thorough evaluation of the accuracy, precision, recall, and trade-off between FPs and FNs of the intrusion detection models. Researchers can assess the efficiency of the suggested framework and its capacity to recognize and counteract cyberattacks in the vehicle's CAN system by examining these metrices.

4. Experimental Results

The RF classifier was applied to predict the cyberattack for using the four datasets with a 70:30 training and testing split of the data. The random state is taken as 11 and the number of estimators is taken as 200. The prediction accuracy for the four datasets (DoS, fuzzy, gear, and RMP) is found to be 0.93, 0.99, 1.0, and 1.0, respectively. The confusion matrix for the DoS dataset is shown in Figure 7.

5. Conclusion and Future Works

In this study, we aimed to develop an effective ML model based on an RF classifier for the intrusion detection of cyberattacks in the CAN of CAVs. The research focused on addressing the vulnerabilities in the CAN protocol, particularly its susceptibility to cyber threats, and proposed a robust ML model for accurate intrusion detection. The high prediction accuracy proves that the RF-based ML model presented in this research is very effective in accurately detecting intrusions in the vehicle CAN of CAVs. Additional complex scenarios of cyberattacks can be studied in future works.

The implications of our research extend beyond the academic realm, with practical applications in enhancing the security of CAVs. The developed ML model offers a reliable means of detecting and mitigating potential cyber threats, thereby safeguarding the communication networks and ensuring the privacy and safety of CAV users. Our work contributes to the ongoing efforts to address the key concerns of safety, privacy, and data security, as identified in public perceptions toward CAVs.

It is crucial to acknowledge the limitations of the proposed method. The model's effectiveness may be influenced by the dynamic nature of cyber threats, and continuous adaptation is necessary to combat evolving intrusion techniques. Additionally, the model's reliance on historical data assumes that future cyber threats will exhibit patterns similar to those observed in the training data.

While the current study marks a significant step in the direction of securing CAVs against cyberattacks, future research should aim at perfecting and expanding the proposed framework. This could involve exploring advanced ML methods, incorporating additional features or datasets, and adapting to emerging attack vectors and tactics. The framework's flexibility and ability to identify and mitigate new risks and vulnerabilities are crucial for its sustained effectiveness. In the practical implementation of the ML model for real-time application, it is acknowledged that inaccuracies, such as FPs or FNs, may arise. Addressing these challenges could involve exploring an active learning paradigm, as suggested in previous works, to minimize label mismatches when dealing with unstructured datasets [29].

Conflicts of Interest

Manoj K. Jha is an Associate Editor for *Journal of Computational and Cognitive Engineering*, and was not involved in the editorial review or the decision to publish this article. The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The CAV datasets data set that support the findings of this study are openly available at http://ocslab.hksecurity.net/Datasets/CAN-intrusion-dataset.

References

[1] Lee, D., & Hess, D. J. (2022). Public concerns and connected and automated vehicles: Safety, privacy, and data security. *Humanities and Social Sciences Communications*, *9*(1), 90. https://doi.org/10.1057/s41599-022-01110-x

[2] Seo, E., Song, H. M., & Kim, H. K. (2018). GIDS: GAN based intrusion detection system for in-vehicle network. In *16th Annual Conference on Privacy, Security and Trust,* 1–6. https://doi.org/10.1109/PST.2018.8514157

[3] Song, H. M., Woo, J., & Kim, H. K. (2020). In-vehicle network intrusion detection using deep convolutional neural network. *Vehicular Communications*, *21*, 100198. https://doi.org/10.1016/j.vehcom.2019.100198

[4] Alkasassbeh, M., & Al-Haj Baddar, S. (2023). Intrusion detection systems: A state-of-the-art taxonomy and survey. *Arabian Journal for Science and Engineering*, *48*(8), 10021–10064. https://doi.org/10.1007/s13369-022-07412-1

[5] Lin, S., Liang, Z., Zhao, S., Dong, M., Guo, H., & Zheng, H. (2024). A comprehensive evaluation of ensemble machine learning in geotechnical stability analysis and explainability. *International Journal of Mechanics and Materials in Design*, *20*, 331–352. https://doi.org/10.1007/s10999-023-09679-0

[6] Pazul, K. (1999). *Controller Area Network (CAN) basics*. Retrieved from: https://cika.com/soporte/Information/Microchip/AnalogInterface/CAN/AppNotes/AN713(DS00713a).pdf

[7] Foster, I., & Koscher, K. (2015). *Exploring controller area networks.* Retrieved from: https://www.usenix.org/system/files/login/articles/login_dec15_02_foster.pdf

[8] Walker, M. (2015). *Security experts reveal how a Tesla model S was hacked.* Retrieved from: https://www.hollywoodreporter.com/news/general-news/security-experts-reveal-how-a-814062/

[9] Kumar, A., & Das, T. K. (2023). CAVIDS: Real time intrusion detection system for connected autonomous vehicles using logical analysis of data. *Vehicular Communications*, *43*, 100652. https://doi.org/10.1016/j.vehcom.2023.100652

[10] Palaniswamy, B., Camtepe, S., Foo, E., & Pieprzyk, J. (2020). An efficient authentication scheme for intra-vehicular controller area network. *IEEE Transactions on Information Forensics and Security*, *15*, 3107–3122. https://doi.org/10.1109/TIFS.2020.2983285

[11] Boudguiga, A., Klaudel, W., Boulanger, A., & Chiron, P. (2016). A simple intrusion detection method for controller area network. In *IEEE International Conference on Communications,* 1–7. https://doi.org/10.1109/ICC.2016.7511098

[12] Keen Security Lab. (2024). *Keen security lab blog.* Retrieved from: https://keenlab.tencent.com/en/tags/CarHacking/

[13] Giannaros, A., Karras, A., Theodorakopoulos, L., Karras, C., Kranias, P., Schizas, N., . . . , & Tsolis, D. (2023). Autonomous vehicles: Sophisticated attacks, safety issues, challenges, open topics, blockchain, and future directions. *Journal of Cybersecurity and Privacy*, *3*(3), 493–543. https://doi.org/10.3390/jcp3030025

[14] Hubballi, N., & Suryanarayanan, V. (2014). False alarm minimization techniques in signature-based intrusion detection systems: A survey. *Computer Communications*, *49*, 1–17. https://doi.org/10.1016/j.comcom.2014.04.012

[15] Jyothsna, V. V. R. P. V., Prasad, R., & Prasad, K. M. (2011). A review of anomaly based intrusion detection systems. *International Journal of Computer Applications*, *28*(7), 26–35.

[16] Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, *2*(3), 160. https://doi.org/10.1007/s42979-021-00592-x

[17] Zheng, Q., Tian, X., Yu, Z., Jiang, N., Elhanashi, A., Saponara, S., & Yu, R. (2023). Application of wavelet-packet transform driven deep learning method in $PM_{2.5}$ concentration prediction: A case study of Qingdao, China. *Sustainable Cities and Society*, *92*, 104486. https://doi.org/10.1016/j.scs.2023.104486

[18] Zheng, Q., Zhao, P., Wang, H., Elhanashi, A., & Saponara, S. (2022). Fine-grained modulation classification using multi-scale radio transformer with dual-channel representation. *IEEE Communications Letters*, *26*(6), 1298–1302. https://doi.org/10.1109/LCOMM.2022.3145647

[19] Zheng, Q., Yang, M., Zhang, Q., & Zhang, X. (2017). Fine-grained image classification based on the combination of artificial features and deep convolutional activation features. In *IEEE/CIC International Conference on Communications in China*, 1–6. https://doi.org/10.1109/ICCChina.2017.8330485

[20] Bari, B. S., Yelamarthi, K., & Ghafoor, S. (2023). Intrusion detection in vehicle Controller Area Network (CAN) bus using machine learning: A comparative performance study. *Sensors*, *23*(7), 3610. https://doi.org/10.3390/s23073610

[21] Basavaraj, D., & Tayeb, S. (2022). Towards a lightweight intrusion detection framework for in-vehicle networks. *Journal of Sensor and Actuator Networks*, *11*(1), 6. https://doi.org/10.3390/jsan11010006

[22] Hossain, M. D., Inoue, H., Ochiai, H., Fall, D., & Kadobayashi, Y. (2020). LSTM-based intrusion detection system for in-vehicle can bus communications. *IEEE Access*, *8*, 185489–185502. https://doi.org/10.1109/ACCESS.2020.3029307

[23] Kalkan, S. C., & Sahingoz, O. K. (2020). In-vehicle intrusion detection system on controller area network with machine learning models. In *11th International Conference on Computing, Communication and Networking Technologies*, 1–6. https://doi.org/10.1109/ICCCNT49239.2020.9225442

[24] Lin, H. C., Wang, P., Chao, K. M., Lin, W. H., & Chen, J. H. (2022). Using deep learning networks to identify cyber attacks on intrusion detection for in-vehicle networks. *Electronics*, *11*(14), 2180. https://doi.org/10.3390/electronics11142180

[25] Alsaade, F. W., & Al-Adhaileh, M. H. (2023). Cyber-attack detection for self-driving vehicle networks using deep autoencoder algorithms. *Sensors*, *23*(8), 4086. https://doi.org/10.3390/s23084086

[26] Nagarajan, J., Mansourian, P., Shahid, M. A., Jaekel, A., Saini, I., Zhang, N., & Kneppers, M. (2023). Machine learning based intrusion detection systems for connected autonomous vehicles: A survey. *Peer-to-Peer Networking and Applications*, *16*(5), 2153–2185. https://doi.org/10.1007/s12083-023-01508-7

[27] Mansourian, P., Zhang, N., Jaekel, A., & Kneppers, M. (2023). Deep learning-based anomaly detection for connected autonomous vehicles using spatiotemporal information. *IEEE Transactions on Intelligent Transportation Systems*, *24*(12), 16006–16017. https://doi.org/10.1109/TITS.2023.3286611

[28] Fang, M., & Zhu, X. (2014). Active learning with uncertain labeling knowledge. *Pattern Recognition Letters*, *43*, 98–108. https://doi.org/10.1016/j.patrec.2013.10.011

[29] Jha, M. K., Wanko, N., & Bachu, A. K. (2023). A machine learning-based active learning framework to capture risk and uncertainty in transportation and construction scheduling. In *Recent Trends in Transportation Infrastructure: Select Proceedings of TIPCE 2022*, *2*, 167–178. https://doi.org/10.1007/978-981-99-2556-8_13

How to Cite: Jha, M. K., & Jaiswal, R. (2024). A Machine Learning Model to Predict Cyberattacks in Connected and Autonomous Vehicles. *Journal of Computational and Cognitive Engineering*, *3*(3), 307–315. https://doi.org/10.47852/bonviewJCCE42022066