# To Design an Intrusion Response System for IIoT

[*]Semester 2 Project

Madhur Patle
*IT department*
IIIT Allahabad
Prayagraj, India
MIT2021009@iiita.ac.in

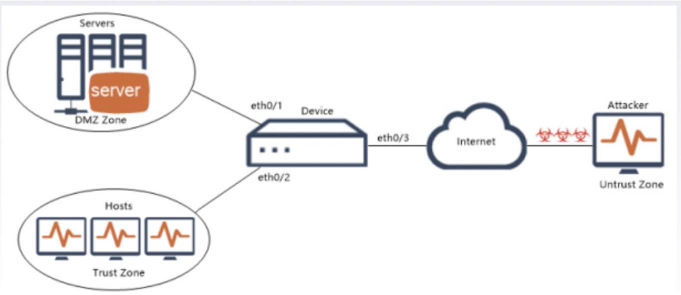
***Abstract*—Our target is to build a intrusion detection system with response component. Autonomous vehicles (AVs) are a potential technology for improving safety and driving efficiency in Intelligent Transportation Systems (ITSs). Vehicle-to-everything (V2X) technology allows vehicles and other infrastructures to communicate. AVs and the Internet of Vehicles (IoV) are, nevertheless, subject to a variety of cyber-attacks, including denial of service, spoofing, and sniffer. An intelligent intrusion detection system (IDS) based on tree-structure machine learning models is proposed in this research. The results of the suggested intrusion detection system's application on common data sets show that the system can detect a variety of cyber-attacks in AV networks. Additionally, the proposed ensemble learning and feature selection methodologies allow the proposed system to attain both a high detection rate and a cheap computing cost and a response component is added which will perform polling to check genuine attacks.**

***Index Terms*—CAN bus, VANET, autonomous vehicles, random forest, XGBoost, stacking, cyber security, intrusion detection system**

## I. Introduction

In the last few decades, humans have become increasingly technology-dependent.The availability of computer networks and the integrity of data must be secure enough from intrusions, which include denial of service (DoS) attacks, unauthorized access, spoofing attacks, and application-layer attacks . The field of intrusion response deals with the problem of once an intrusion is detected. How can the system be protected? Its goal is to handle the attack in such a way that damage is minimized. IDS is activated when some intrusions are detected in the system. IRS is always activated on the basis of IDS output. When IDSs obtain threat information, the response component generates responses on the basis of the symptoms of attacks.

The traditional vehicular ad hoc networks (VANETs) are rapidly transforming into the Internet of Vehicles (IoV) as more vehicles, devices, and infrastructures get involved [1]. VANETs enable wireless communications between cars and devices in intelligent transportation systems (ITSs), then transform the vehicles and gadgets into wireless routers or mobile nodes [2]. Autonomous vehicles (AV) are a rapidly developing technology that offers a viable option for reducing traffic collisions and associated costs. Vehicle-to-everything (V2X) [2] technology enables local and wide-area cellular network connections between automobiles, pedestrians, and infrastructures. With wireless communications, V2X technology intends to connect more Internet of things (IoT) devices. Some of these devices, however, lack security features like firewalls and gateways [3]. Because assaulting or maliciously managing automobiles on the road poses a substantial hazard to human lives, AVs are vulnerable to network threats with catastrophic repercussions. The following attacks are examples of potential networking threats. Denial of service (DoS) assaults, which transmit a huge number of irrelevant messages or requests to occupy a node, are a popular sort of attack. The attackers perform spoofing attacks such as GPS spoofing to masquerade as legitimate users and provide the nodes with fake GPS information [4]. Sniffing attacks such as port scan attacks is another type of attack, which is launched to obtain confidential or sensitive data of the vehicles systems and users [5].



## II. Problem definition

An intelligent intrusion detection system (IDS) is proposed based on tree-structure machine learning models. The results from the implementation of the proposed intrusion detection system on standard data sets indicate that the system has the ability to identify various cyber-attacks in the AV networks. Furthermore, the proposed ensemble learning and feature selection approaches enable the proposed system to achieve high detection rate and low computational cost simultaneously

## III. Literature Review

| S.no. | Paper | Summary | Limitations |
|---|---|---|---|
| 1 | Tree-based Intelligent Intrusion Detection System in Internet of Vehicles | Proposed intrusion detection system on standard data sets indicate that the system has the ability to identify various cyber-attacks in the AV networks. | All model accuracy can be combine with stacking |
| 2 | Intrusion response systems: Foundations, design, and challenges | Types of IRS, Design Parameters, Characteristics and Challenges to building intrusion response system. | IRS lack efficient algorithms to update the response history over time. |
| 3 | An Intrusion Response Approach for Elastic Applications Based on a Reinforcement Learning | IRS based on deep reinforcement learning and transfer learning, which automatically adapts to system changes | Approach should support distributed ensemble of IRSs, in order to possibly reduce the training time |
| 4 | Parti, Rohit. (2004). Design of an Intrusion Response System using Evolutionary Computation | Presents an evolutionary computation method to tackle the challenges to design an automated intrusion response system. | A response model generated only on previously occurred intrusions. |
| 5 | InDReS: An Intrusion Detection and Response System for the Internet of Things with 6LoW AN | InDReS is developed for resource-constrained loT network for detecting sinkhole attacks, using constrained based specification technique | Use behavioral rule based specification and considering design parameters |

## IV. Dataset Details

Anomaly-based Network Intrusion Detection Systems (NIDSs) are designed to learn and extract complicated network data behaviours in order to categorise incoming traffic as malicious or benign. Various features sent through network traffic, such as packet counts/sizes, protocols, services, and flags, can be used to generate network attack vectors.

### A. *ToN - IoT*

ToN-IoT- A recent heterogeneous dataset launched in 2020 [9] that contains telemetry data from Internet of Things (IoT) services, network traffic from IoT networks, and the component of the operating system log including network traffic flows. A significant number of records make up the dataset. Various attack scenarios carried out in a realistic medium-scale simulation ACCS's network at the Cyber Range Lab. Bro-IDS is currently known as Zeek. The dataset's original 44 features were extracted using this method. The dataset consists of There were 796,380 benign flows (3.56 percent) and 21,542,641 assault samples (96.44 percent) totaling 22,339,021 flows.

| Class | Count | Summary |
|---|---|---|
| Benign | 270279 | Normal unmalicious flows |
| Backdoor | 17247 | A technique that aims to attack remote-access computers by replying to specific constructed |
| DoS | 17717 | An attempt to overload a computer system's resources with the aim of preventing access to or availability of its data client applications |
| DDoS | 326345 | An attempt similar to DoS but has multiple different distributed sources. |
| Injection | 468539 | A variety of attacks that supply untrusted inputs |
| MITM | 1295 | Man In The Middle is a method that places an attacker between a victim and host |
| Password | 156299 | covers a variety of attacks |
| Ransomware | 142 | An attack that encrypts the files stored on a host |
| Scanning | 21467 | A group that consists of a variety of techniques |
| XSS | 99944 | Cross-site Scripting is a type of injection |

### B. *CIC-IDS2018 datasets*

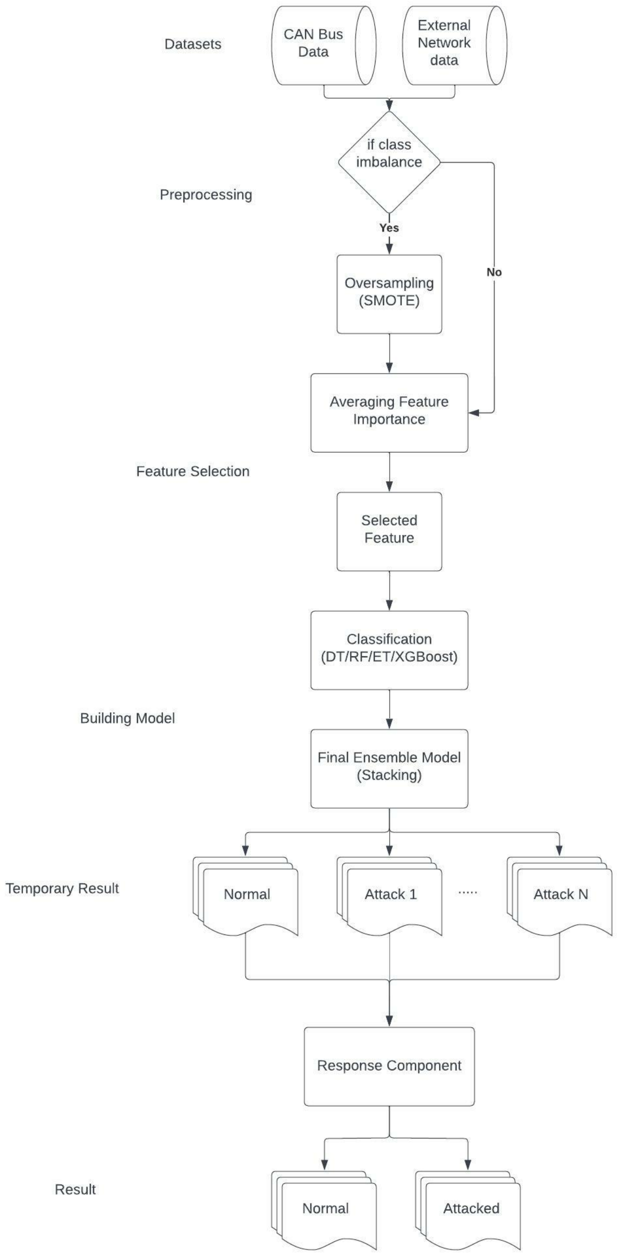
In 2018, the Communications Security Establishment (CSE) and the Canadian Institute for Cybersecurity (CIC) collaborated on a dataset called CSE-CIC-IDS2018 [14]. The victim network was created in the year. Five separate organisational departments and an additional server room are depicted in a realistic manner. Realistic network traffic generated the benign packets. occurrences based on abstract human user behaviour The possible attack scenarios were carried out by one or more machines not connected to the target network. The 75 features in the original dataset were extracted using the

CICFlowMeter-V3 programme. There are 13,484,708 benign flows (83.07 percent) and 2,748,235 harmful flows in the entire dataset. There were 16,232,943 attack flows (16.93 percent) in total.

| Class | Count | Summary |
|---|---|---|
| Benign | 7373198 | Normal unmalicious flows |
| BruteForce | 287597 | A technique that aims to obtain usernames and password credentials by accessing a list of predefined possibilities |
| Bot | 15683 | An attack that enables an attacker to remotely control several hijacked computers to perform malicious activities |
| DoS | 269361 | An attempt to overload a computer system's resources with the aim of preventing access to or availability of its data. |
| DDoS | 380096 | An attempt similar to DoS but has multiple different distributed sources. |
| Web Attacks | 4394 | A group that includes SQL injections, command injections and unrestricted file uploads |
| Infiltration | 62072 | An inside attack that sends a malicious file |

## V. Project Flow

Step by step implementation of proposed methodology The suggested IDS is deployed in numerous areas inside the AV system to provide protection for both intra-vehicle and external communications. Detecting risks on the internet. Secure the May bus, and the IDS can be installed on top. Every sent message is processed by the CAN bus. Ensure that the nodes are secure [9]. In addition, the To secure the gateway, the proposed IDS can be installed inside. The proposed model's procedure is as follows. First, enough network traffic data is gathered. Second, if the data set's classes are unbalanced, oversampling is used to mitigate the effect. To reduce computational cost, the following stage involves feature selection based on averaging feature importance. Following that, four base-models are created as inputs to the stacking ensemble model. Finally, a model is created to classify the data. Because networks in real life retain a normal state most of the time, network data is frequently class unbalanced, and attack-label instances are frequently insufficient. Random oversampling and Synthetic Minority Oversampling Technique (SMOTE) can be used to generate extra data in the minority classes that do not have enough data to solve the issue of class-imbalanced data, which often results in a low anomaly detection rate. The basic method of random oversampling is to simply duplicate the samples in order to boost sample sizes in minority groups.



### A. The proposed ML approaches

The development of the IDS in the proposed system can be thought of as a multi-classification problem, and machine learning methods are commonly utilised to address such classification problems [12] [13]. Decision tree, random forest, additional trees, and XGBoost are some of the ML techniques that are based on tree structure. A common classification method based on the divide and conquer strategy is the decision tree (DT) [14]. A decision tree (DT) is made up of decision nodes and leaf nodes, which indicate a judgement test over one of the features and the result class, respectively. Random forest (RF) [15] is a classifier for ensemble learning. According to the majority voting rule, the class with the most votes wins. Decision trees were chosen as the classification method.  result. Extra trees (ET) [16] is another ensemble method.  a model built from a set of randomised decision tees Created by combining portions of the data set In XGBoost [17], on the other hand, is an ensemble learning method. developed to boost speed and performance through the use of To integrate several decision trees, use the gradient descent approach.

## VI. Result

Performance evaluation of all models

| Method | Method Acc% | DR % | FAR % | F1 Score | Execution Time (S) |
|---|---|---|---|---|---|
| KNN | 97.4 | 96.3 | 5.3 | 0.934 | 911.6 |
| SVM | 96.5 | 95.7 | 4.8 | 0.933 | 13765.6 |
| DT | 99.99 | 99.99 | 0.006 | 0.999 | 328 |
| RF | 99.99 | 99.99 | 0.0003 | 0.999 | 506.8 |
| ET | 99.99 | 99.99 | 0.0005 | 0.999 | 216.3 |
| XGBoost | 99.98 | 99.98 | 0.012 | 0.999 | 3499.1 |
| Stacking | 99.99 | 99.99 | 0.0006 | 0.999 | 325.6 |

The trials were run on a machine with a 6 Core i7-8700 processor and 16 GB of memory, and the proposed system was constructed using Python 3.5. Tables III and IV illustrate the results of testing several algorithms on the CAN-intrusion data set and the CICIDS2017 data set, respectively. Table III shows that the proposed system's tree-based algorithms, comprising DT, RF, ET, and XGBoost, are 2.5 percent more accurate than KNN and 3.4 percent more accurate than SVM when tested on the CAN-intrusion data set.

In addition, DT, RF, ET, and XGBoost provide multi-threading, leading in faster execution times than KNN and SVM. Only the other three algorithms, DT, RF, and ET, were selected into the stacking ensemble model, and the single model with the best performance, RF, was selected to be the meta-classifier in the second layer, because XGBoost has the lowest accuracy and longest execution time among the four tree-based ML models. After combining the three tree-based models with stacking, the accuracy, detection rate, and F1 score all reach 100

## References

[1] Z. Inayat, A. Gani, N. B. Anuar, M. K. Khan, and S. Anwar, "Intrusion response systems: Foundations, design, and challenges," J. Netw. Comput. Appl., vol. 62, pp. 53–74, 2016

[2] S. Iannucci, E. Casalicchio, and M. Lucantonio, "An intrusion response approach for elastic applications based on reinforcement learning," in 2021 IEEE Symposium Series on Computational Intelligence (SSCI), 2021

[3] Parti, Rohit. (2004). Design of an Intrusion Response System using Evolutionary Computation

[4] Ghumro, Adnan  Kanwal, Aisha  Memon, Irfana  Simming, Insaf. (2021). A Review of Mitigation of Attacks in IoT using Deep Learning Models. 18. 36-42

[5] Arnaboldi, Luca  Morisset, Charles. (2021). A Review of Intrusion Detection Systems and Their Evaluation in the IoT

[6] Yang, L., Moubayed, A., Hamieh, I.,  Shami, A. (2019). Tree-based intelligent intrusion detection system in internet of vehicles. 2019 IEEE Global Communications Conference (GLOBECOM).

[7] Sarhan, M., Layeghy, S., Moustafa, N.,  Portmann, M. (2021). NetFlow datasets for machine learning-based network intrusion detection systems. In Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (pp. 117–135). Springer International Publishing.