21AIE433 IPS-IDS
PROJECT PRESENTATION

A MULTI-TIERED HYBRID INTRUSION DETECTION SYSTEM FOR INTERNET OF VEHICLES

Guided by
Niranjan DK

Team Infinity

APOORVA M	BL.EN.U4AIE19007
TANUJ M	BL.EN.U4AIE19041
AISHWARYA V	BL.EN.U4AIE19068

ABSTRACT

- Modern vehicles are connected to external networks through vehicle-to-everything technologies, enabling their communications with other vehicles, infrastructures, and smart devices.
- However, the improving functionality and connectivity of modern vehicles also increase their vulnerabilities to cyber-attacks targeting both intra-vehicle and external networks due to the large attack surfaces.
- To secure vehicular networks, many researchers have focused on developing intrusion detection systems (IDSs) that capitalize on machine learning methods to detect malicious cyberattacks.
- A multi-tiered hybrid IDS that incorporates a signature-based IDS and an anomaly-based IDS is proposed to detect both known and unknown attacks on vehicular networks.

INTERNET OF VEHICLES

- With the increasing research and rapid development of the Internet of Vehicles (IoV) technology, connected vehicles (CVs) and autonomous vehicles (AVs) are becoming increasingly popular in the modern world .

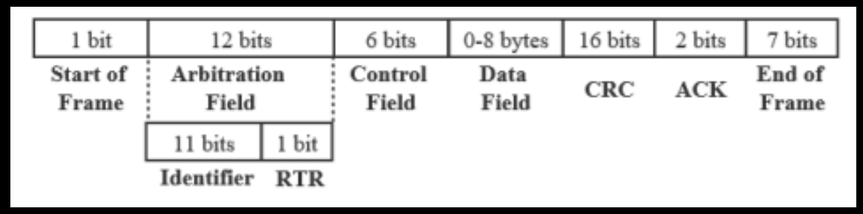
- IoV serves as a primary vehicular communication framework that enables reliable communications between vehicles and other IoV entities, such as infrastructures, pedestrians, and smart devices.

INTRA-VEHICLE NETWORKS

- Vehicles today are equipped with more and more sensors.
- IVNs involve an increasing number of electronic control units (ECUs) to adopt various functionalities .
- All ECUs in a vehicle are connected by a controller area network (CAN) bus to transmit messages and perform actions.

VULNERABILITIES

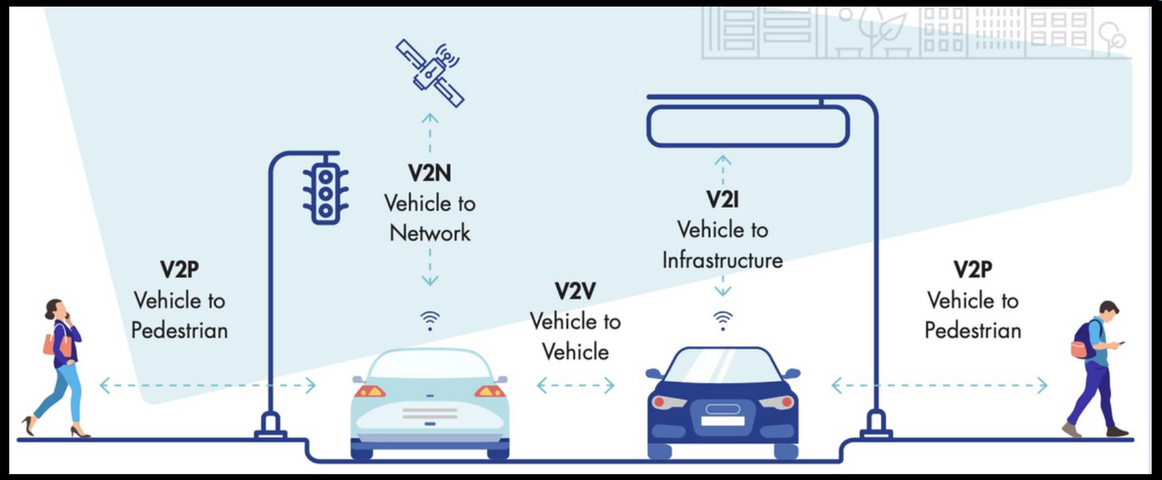
- CAN is a bus communication protocol that defines an international standard for efficient and reliable intra-vehicle communications among ECUs
- CAN is the most common type of IVN due to its low cost and complexity, high reliability, noise resistance, and fault-tolerance properties
- However, CAN is vulnerable to various cyber threats due to its broadcast transmission strategy, lack of authentication and encryption, and unsecured priority scheme



- Among all fields, the data field with the size of 0-8 bytes is the most important and vulnerable one, since it contains the actual transmitted data that determines the node actions
- Message injection attacks are the primary type of intravehicle attack
- it can further be classified as Dos Attakcs, Fuzzy attacks and spooking attacks by their objective

VEHICLE TO EVERYTHING

- external networks connect modern vehicles to the outer environment by vehicleto-everything (V2X) technologies
- V2X technology allows modern vehicles to communicate with other vehicles, roadside infrastructures, and road users

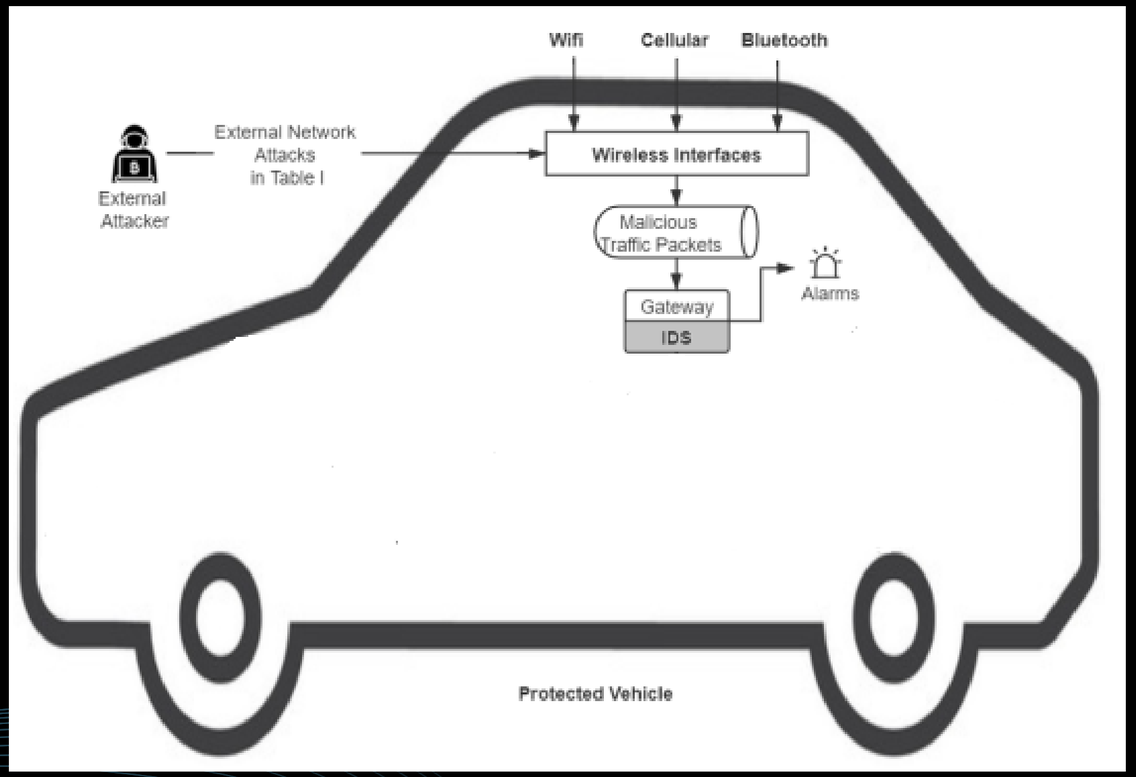


VULNERABILITIES

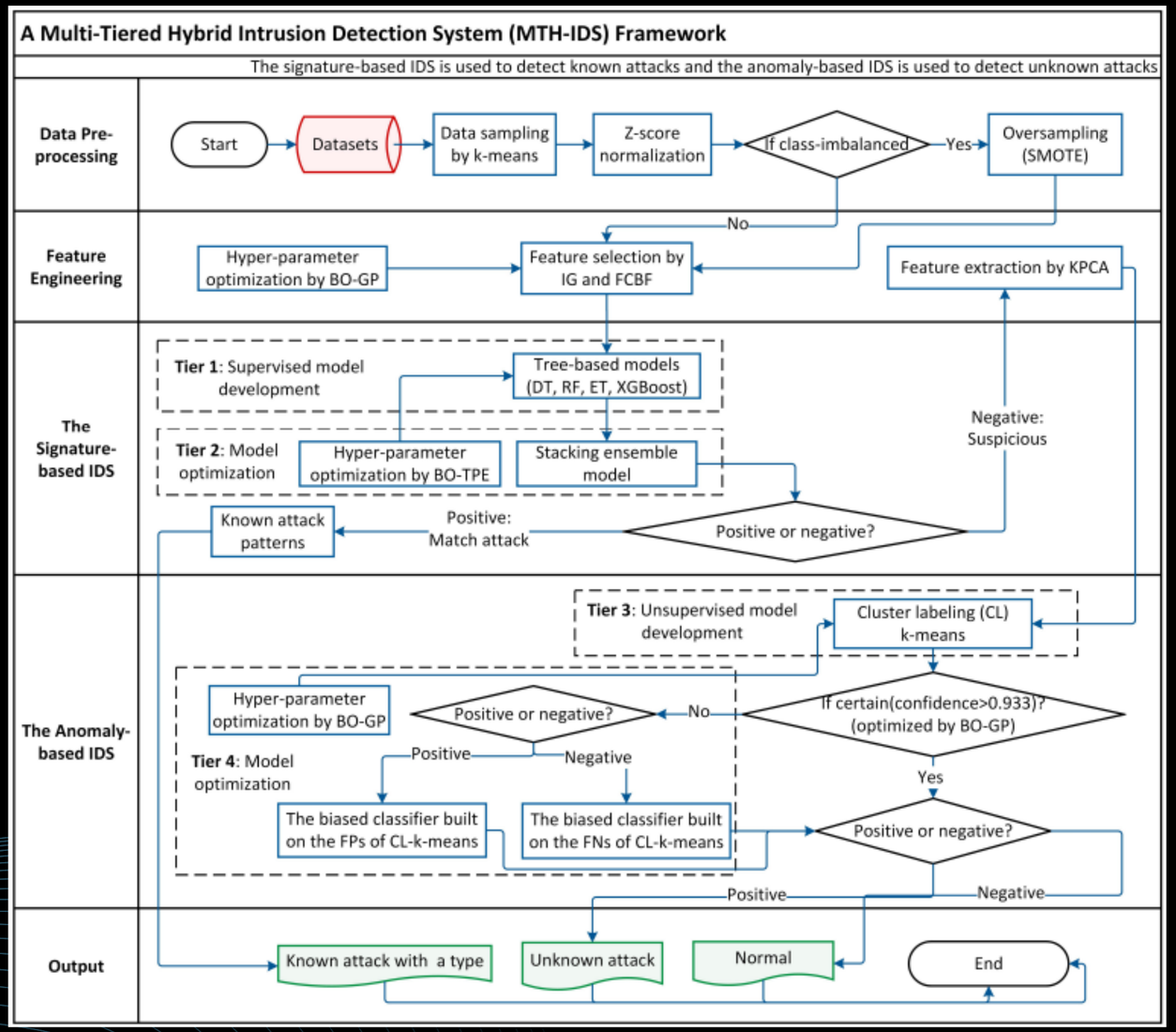
- With the increasing connectivity of modern IoV, external vehicular networks are becoming large networks that involve various other networks and devices.
- Thus, external vehicular networks are vulnerable to various general cyber threats because each vehicle or device is a potential entry point for intrusions
- Typical attacks in IoV include DoS, GPS spoofing, jamming, sniffing, brute-force, Botnets, infiltration, and web attacks

Attack Type	Description and IoV Scenarios
DoS [32]	Send a large number of requests to exhaust the compromised nodes' resources, causing vehicle unavailability or accidents.
GPS Spoofing [32]	Masquerade as authorized IoV users to provide a node with false information, like false geographic information, therefore causing fake evidence, event delay, or property losses.
Jamming [32]	Jam signals to prevent legitimate IoV devices from communicating with connected vehicles.
Sniffing [32]	Capture vehicular network packets to steal confidential or sensitive information of vehicles, users, or enterprises.
Brute-force [11]	Crack passwords in vehicle systems to take control of vehicles or machines and perform malicious actions.
Botnets [11]	Infect multiple connected vehicles and IoV devices with Bot viruses to breach them and launch other attacks.
Infiltration [11]	Traverse the compromised vehicle systems and create a backdoor for future attacks.
Web Attack [11]	Hack IoV servers or web interfaces of connected vehicles to gain confidential information or perform malicious actions.

PROPOSED IDS-PROTECTED VEHICLE ARCHITECTURE



FRAMEWORK OF PROPOSED MTH-IDS



SUPERVISED MODELS - Signature based Detection

- **DECISION TREE** - In this technique, we split the population or sample into two or more sub-populations based on most significant splitter / differentiator in input variables.

- **RANDOM FOREST** - Random forest is an ensemble of many decision trees. Random forests are built using a method called bagging in which each decision trees are used as parallel estimators.

- **EXTRA TREES** - Very similar to a Random Forest Classifier and only differs from it in the manner of construction of the decision trees in the forest.

- **XGBOOST** - XGBoost, which stands for Extreme Gradient Boosting, is a scalable, distributed gradient-boosted decision tree (GBDT) machine learning library. It provides parallel tree boosting

UNSUPERVISED MODELS - Anomaly based Detection

K- MEANS CLUTSERING

- It allows us to cluster the data into different groups and a convenient way to discover the categories of groups in the unlabeled dataset on its own without the need for any training.

- It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters.

HYPER-PARAMETER OPTIMIZATION METHODS

Bayesian optimization with tree-based Parzen estimator (BO-TPE)

- The default hyper-parameters of ML algorithms often cannot return the best model. BO-TPE can optimize the models' hyper-parameters to obtain the optimized base classifiers.

Bayesian optimization with Gaussian processes (BO-GP)

- CL-k-means has an important hyper-parameter, the number of clusters, k. BO-GP is an effective HPO method to optimize k and obtain the optimized CL-k-means model

DATA SET - CICID

- Is a reliable as it is a representative dataset of current external networks because it is the most state-of-the-art dataset and contains more features, instances, and cyber-attack types than other datasets.
- For minority classes with small numbers of samples (from 36 to 13,835), the SMOTE method was implemented to synthesize more samples to enable the minority classes to have at least 100,000 samples.
- Addressing class-imbalance can avoid obtaining biased models with low attack detection rates.

Class Label	Corresponding Attack Type in Table II [78]	Original Number of Samples	Number of Training Set Samples After Balancing	Number of Test Set Samples
BENIGN	-	2,273,097	1,591,168	681,929
Bot	Botnets	1,966	100,000	590
DDoS	DoS	380,699	266,489	114,210
DoS GoldenEye				
DoS Hulk				
DoS Slow-httptest				
DoS Slowloris				
Heartbleed				
Port-Scan	Sniffing	158,930	111,251	47,679
SSH-Patator	Brute-Force	13,835	100,000	4,150
FTP-Patator				
Infiltration	Infiltration	36	100,000	11
Web Attack – Brute Force	Web Attack	2,180	100,000	654
Web Attack – Sql Injection				
Web Attack – XSS				

RESULTS

Signature based IDS Results

Accuracy of XGBoost: 0.9957089552238806
Precision of XGBoost: 0.9956893766598436
Recall of XGBoost: 0.9957089552238806
F1-score of XGBoost: 0.9956902750637269

XGBoost

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3645
1	0.99	1.00	1.00	393
2	1.00	1.00	1.00	19
3	1.00	1.00	1.00	609
4	0.83	0.71	0.77	7
5	0.99	1.00	0.99	251
6	0.99	0.99	0.99	436
accuracy			1.00	5360
macro avg	0.97	0.96	0.96	5360
weighted avg	1.00	1.00	1.00	5360

Random Forest

Accuracy of RF: 0.9951492537313433
Precision of RF: 0.9951646455154706
Recall of RF: 0.9951492537313433
F1-score of RF: 0.9951217831414103

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3645
1	0.99	1.00	0.99	393
2	1.00	1.00	1.00	19
3	1.00	1.00	1.00	609
4	1.00	0.71	0.83	7
5	0.98	1.00	0.99	251
6	0.99	0.99	0.99	436
accuracy			1.00	5360
macro avg	0.99	0.96	0.97	5360
weighted avg	1.00	1.00	1.00	5360

Accuracy of ET: 0.9955223880597015
Precision of ET: 0.9955353802920419
Recall of ET: 0.9955223880597015
F1-score of ET: 0.9954932494250629

Extra Trees

	precision	recall	f1-score	support
0	1.00	1.00	1.00	3645
1	0.99	1.00	1.00	393
2	1.00	1.00	1.00	19
3	0.99	1.00	0.99	609
4	1.00	0.71	0.83	7
5	0.98	1.00	0.99	251
6	1.00	0.99	0.99	436
accuracy			1.00	5360
macro avg	0.99	0.96	0.97	5360
weighted avg	1.00	1.00	1.00	5360

Decision Trees

Accuracy of DT: 0.9936567164179104
Precision of DT: 0.9940667447648622
Recall of DT: 0.9936567164179104
F1-score of DT: 0.9938179408993949

	precision	recall	f1-score	support
0	1.00	0.99	1.00	3645
1	0.99	1.00	0.99	393
2	1.00	1.00	1.00	19
3	1.00	1.00	1.00	609
4	0.45	0.71	0.56	7
5	0.98	1.00	0.99	251
6	0.99	0.98	0.98	436
accuracy			0.99	5360
macro avg	0.92	0.95	0.93	5360
weighted avg	0.99	0.99	0.99	5360

Result after Stacking of all ML Algorithms

```
Accuracy of XGBoost: 0.9957089552238806
Precision of XGBoost: 0.9956893766598436
Recall of XGBoost: 0.9957089552238806
F1-score of XGBoost: 0.9956902750637269
              precision    recall  f1-score   support

           0       1.00      1.00      1.00      3645
           1       0.99      1.00      1.00       393
           2       1.00      1.00      1.00        19
           3       1.00      1.00      1.00       609
           4       0.83      0.71      0.77         7
           5       0.99      1.00      0.99       251
           6       0.99      0.99      0.99       436

    accuracy                           1.00      5360
   macro avg       0.97      0.96      0.96      5360
weighted avg       1.00      1.00      1.00      5360
```

Anomaly based IDS Result

	precision	recall	f1-score	support
0	0.99	0.90	0.94	1255
1	0.91	0.99	0.95	1255
accuracy			0.95	2510
macro avg	0.95	0.95	0.94	2510
weighted avg	0.95	0.95	0.94	2510

```
0.9450199203187251
[[1127  128]
 [  10 1245]]
```

CONCLUSION

- To enhance IoV security, this work proposed a multi-tiered hybrid intrusion detection system (MTH-IDS) model that can detect various types of known and zero-day cyber-attacks on external-vehicular networks for modern vehicles.
- Through data pre-processing and feature engineering, the quality of the input data can be significantly improved for more accurate model learning.
- The proposed system can effectively detect various types of known attacks with an accuracy of 99.88% and unknown attacks with average F1-scores of 0.800 on the CICIDS2017 dataset
- The experimental results on a vehicle-level machine also show the feasibility of the proposed system in real-time environments

FUTUREWORK

- The performance of the model for known attacks was very immpressive. Even though the perfomance of the models for zero day attacks was appreciable, there is a scope for improvement.
- Hence in future work, the proposed anomaly-based IDS framework) can be further improved by doing research on other unsupervised learning and online learning methods.
- This work concentrates on detecting attacks on Vehicle-to -Everything network, similarly models for detecting attacks on intra-vehicle networks can be included

REFERENCES

1. H. Liang et al., “Network and system level security in connected vehicle applications,” IEEE/ACM Int. Conf. Comput. Des. Dig. Tech. Pap. ICCAD, pp. 1–7, 2018.
2. M. Gmiden, M. H. Gmiden, and H. Trabelsi, “An intrusion detection method for securing in-vehicle CAN bus,” 2016 17th Int. Conf. Sci. Tech. Autom. Control Comput. Eng. STA 2016 - Proc., pp. 176–180, 2017.
3. J. Liu, S. Zhang, W. Sun, and Y. Shi, “In-vehicle network attacks and countermeasures: Challenges and future directions,” IEEE Netw., vol. 31, no. 5, pp. 50–58, 2017
4. O. Y. Al-Jarrah, C. Maple, M. Dianati, D. Oxtoby, and A. Mouzakitis, “Intrusion Detection Systems for Intra-Vehicle Networks: A Review,” IEEE Access, vol. 7, pp. 21266–21289, 2019
5. L. Yang, ”Comprehensive Visibility Indicator Algorithm for Adaptable Speed Limit Control in Intelligent Transportation Systems”, M.A.Sc. thesis, University of Guelph, 2018.
6. J. Golson, “Jeep hackers at it again, this time taking control of steering and braking systems,” The Verge, Aug. 2016. [Online]. Available: https://www.theverge.com/2016/8/2/12353186/car-hackjeep-cherokee-vulnerability-miller-valasek. [Accessed: 11-Nov-2020].
7. L. Yang, A. Moubayed, I. Hamieh, and A. Shami, “Tree-based Intelligent Intrusion Detection System in Internet of Vehicles,” proc. 2019 IEEE Glob. Commun. Conf., pp. 1–6, Hawaii, USA, 2019.
8. Q. Wang, Y. Qian, Z. Lu, Y. Shoukry, and G. Qu, “A delay based plug-inmonitor for Intrusion Detection in Controller Area Network,” Proc. 2018 Asian Hardw. Oriented Secur. Trust Symp. AsianHOST 2018, pp. 86–91, 2019.

AMRITA
VISHWA VIDYAPEETHAM

THANK YOU