Annals of Mathematics and Computer Science ISSN: 2789-7206
Vol 35 (2026) 1-9 https://doi.org/10.56947/amcs.v35.871

EVALUATING ZERO-DAY GENERALISATION IN VANET DETECTION

ANTHONY IBRAHIM

Abstract. The rapid expansion of the Internet of Vehicles has introduced security vulnerabilities, requiring robust Intrusion Detection Systems. This study evaluates the generalisation capabilities of Random Forest, XGBoost, Naive Bayes, Logistic Regression and Extra Trees against zero-day vehicular attacks. Using the VeReMi NextGen dataset, the methodology separates known and unseen attacks using the Leave-One-Attack-Out strategy to simulate realistic, unseen threat scenarios. Results demonstrate that while Random Forest excels at identifying known threats, it suffers from overfitting, resulting in a low F1-score for zero-day attacks. Conversely, XGBoost and Extra Trees exhibit more consistent zero-day performance, suggesting improved robustness against novel vehicular attacks.

Keywords. Internet of Vehicles (IoV), Intrusion Detection, Machine Learning, Zero-day Attacks

1. Introduction

The advent of the Internet of Vehicles (IoV) has revolutionised intelligent transportation systems by facilitating real-time communication between vehicles and infrastructure. However, this hyperconnectivity significantly expands the attack surface, exposing vehicular networks to sophisticated cyber threats ranging from message falsification to denial-of-service attacks. As highlighted by Ferrag et al. [5], integrating deep learning with standard machine learning approaches is becoming essential to securing these complex environments against increasingly malicious intrusions.

Current Intrusion Detection Systems (IDS) often rely on supervised learning models trained on historical data. While these models achieve high accuracy on known patterns, they frequently struggle with "zero-day" attacks—threats that were not present in the training set. Ghaleb et al. [1] emphasise that while ensemble, adaptive and anomaly-based VANET IDS approaches can improve detection performance [2, 16], the ability of a model to generalise its learning to unseen attack vectors remains a critical hurdle. Thus, the results of this study

Date: Received: Jun 1, 2026; Revised: Jul 1, 2026; Accepted: Jul 8, 2026.
* Corresponding author

have several implications in computer science applications such as cybersecurity, anomaly detection and robust machine learning evaluation. The LOAO evaluation framework can aid researchers to test their models on unseen attack types before deployment. Such implementation is useful for designing safer vehicular networks that remain reliable in real world conditions.

This paper addresses the generalisation challenge by evaluating machine learning models against unseen vehicular attacks using a Leave-One-Attack-Out (LOAO) strategy. We utilise the VeReMi NextGen dataset to evaluate five distinct algorithms: Random Forest, XGBoost, Naive Bayes, Logistic Regression and Extra Trees. By separating known and unseen attacks, this study evaluates model generalisation under zero-day conditions. Unlike conventional benchmark-driven VANET IDS studies, this work evaluates model robustness to unseen attacks using a Leave-One-Attack-Out framework. This setup better reflects real-world deployment scenarios where IDS models encounter previously unseen malicious behaviours. Therefore, this study emphasises benchmark-to-zero-day generalisation and robustness rather than benchmark accuracy alone.

The main contributions of this study are a LOAO-based zero-day evaluation of ML-based VANET intrusion detection using the VeReMi NextGen dataset. It compares benchmark and zero-day attack performance, showing that high benchmark accuracy does not guarantee robust generalisation. In addition, Welch's t-test is used to statistically compare the benchmark and the mean zero-day F1-score.

2. Methodology

2.1. **Dataset.** This study is based on the VeReMi NextGen dataset [6] based on the Highway 2 scenario. The Highway 2 scenario was selected for its balanced traffic complexity and diverse vehicular communication behaviours, which are suitable for controlled zero-day evaluation. The dataset includes features such as position, speed, acceleration, heading, timing, and communication noise variables. Additional engineered features were derived from the original sender and receiver vehicular attributes to capture communication consistency and abnormal behavioural patterns. Inspired by the preprocessing framework proposed in [15], absolute positional, speed, and acceleration noise features, as well as sender–receiver relational features such as speed difference, acceleration difference, and Euclidean positional distance between communicating vehicles were generated. The engineered features were defined as speed difference between sender and receiver $|v_s - v_r|$, acceleration difference as $|a_s - a_r|$, and Euclidean positional distance as $\sqrt{(x_s - x_r)^2 + (y_s - y_r)^2}$. Absolute noise features were also used for position, speed, and acceleration to represent the magnitude of measurement disturbance. These engineered features were designed to improve the detection of anomalous vehicular behaviour and to enhance generalisation against zero-day attacks. Additionally, since this paper aims to detect attacks rather than classify them, each attack type was converted into a binary classification (0 = benign, 1 = attack).

TABLE 1. VeReMi NextGen Dataset Summary

Property	**Value**
Samples	1,733,541
Classes	15 (including benign)
Benign	1,370,938
Attack	362,603
Zero-Day Attacks	5

Initially, the dataset was tested using all attack types. This was done to serve as the benchmark for the zero-day experiment. To test the zero-day attacks, a Leave-One-Attack-Out (LOAO) strategy was applied, in which the models were trained on all attack types except the one being tested. Let $A = \{A_1, A_2, \ldots, A_K\}$ denote the set of attack types. For each LOAO iteration i, the model is trained on $A \setminus \{A_i\}$ and tested on A_i.

The original Train and Validation splits were merged to maximise the amount of data available for model learning, while the Test split remained completely unseen during training and was used only for final evaluation. The five attacks chosen for testing are data replay, DoS attack, reversed heading, sudden stop, and traffic congestion sybil. These attacks were selected to represent diverse vehicular attack behaviours and communication anomalies within VANET environments.

The implementation and code can be found on GitHub: https://github.com/anthonyibr24/Zero-Day-attack-on-VeReMi-dataset.

2.2. **ML and Metrics.** This study utilised five machine learning (ML) models, namely: Random Forest (RF), XGBoost, Naive Bayes, Logistic Regression (LR) and Extra Trees. These models were chosen to represent different detection algorithms. For instance, RF builds numerous independent decision trees, whereas XGBoost builds them sequentially, with each tree correcting previous errors. The Naive Bayes model was used because it follows a probabilistic approach based on Bayes' theorem, which assumes independent features [11], while LR was included as a linear baseline classifier. Lastly, Extra Trees (ET) was included for its robustness to noisy tabular data and its ability to generalise across complex behavioural patterns in zero-day attack detection [7].

To test these models, various metrics were used. A 4-fold cross-validation was done by combining the train and validation sets and testing on the test sets. Accuracy measures overall prediction correctness, while precision indicates how reliable a model is when predicting a positive class, and recall measures the proportion of attacks successfully detected. Along with these metrics, ROC-AUC showcases how well the model can differentiate between normal and attack messages. To further evaluate model robustness, Welch's t-test ($\alpha = 0.05$) was applied to compare the benchmark and zero-day average F1-scores across all attacks for each model. The null and alternative hypotheses for Welch's t-test are defined as: $H_0 : \mu_{\text{benchmark}} = \mu_{\text{zero-day}}$ and $H_a : \mu_{\text{benchmark}} \neq \mu_{\text{zero-day}}$. The Welch's t-test used $n_b = 4$, representing the four benchmark cross-validation folds and $n_z = 5$ representing five F1-scores from the five LOAO unseen attacks,

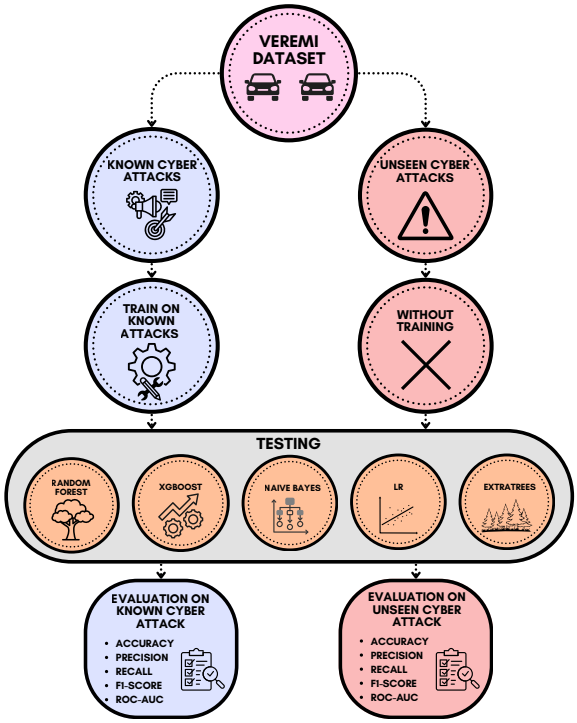

FIGURE 1. Zero-Day Detection Pipeline where models are trained on known cyber-attacks and tested on unseen cyber-attacks using different ML models

which are used to compute the zero-day mean. Hence, the degrees of freedom (df) were estimated using the Welch–Satterthwaite equation, resulting in $df \approx 4$.

Median imputation was applied within each model pipeline, with standardisation used for Logistic Regression and Naive Bayes. A stratified 4-fold cross-validation with shuffling and a fixed random seed of 42 was used for reproducibility. Class imbalance was addressed using class-weighted training for RF, LR and Extra Trees, and `scale_pos_weight` for XGBoost.

3. RESULTS

This section presents the ML results obtained from the benchmark and the zero-day attacks. Referring to Table 3, although all five models appear competitive in terms of accuracy, the F1-scores indicate that Naive Bayes and LR performed poorly. This observation is important, as accuracy alone can be misleading in malicious classification problems due to class imbalance. With RF, XGBoost, and Extra Trees achieving F1-scores of 0.65, 0.70, and 0.69, respectively, these results demonstrate comparatively stronger detection capability. Conversely, Naive Bayes and LR achieved F1-scores of 0.15 and 0.33, respectively, highlighting the dataset's complexity and the limitations of linear and probabilistic approaches. Moreover, XGBoost achieved the highest recall, 0.77, followed by RF (0.71)

and Extra Trees (0.68), demonstrating these models' ability to correctly identify cyber-attacks. However, in terms of precision, RF had a high recall (0.71) but low precision (0.59). This result suggests that RF misclassified numerous benign cases as malicious, resulting in a drop in precision. Therefore, unlike Extra Trees that had consistent precision, recall, and F1-score, RF had a higher rate of false alarms.

Table 4 shows the LOAO zero-day attacks. Overall, the results highlight the challenges of zero-day intrusion detection in VANET environments. This is evident from the significant drop in F1-scores compared to the benchmark experiment. Although several models maintained high recall values, with values up to 1.00, precision remained consistently low for most unseen attacks. This suggests that the models produced a high number of false-positive detections, with some models tending to classify most samples as malicious. This may be influenced by the imbalanced attack ratios. Among the evaluated models, XGBoost and Extra Trees achieved the highest overall F1-scores. In particular, both models achieved the highest F1-scores for the traffic congestion sybil attack, with XGBoost achieving 0.47 and Extra Trees achieving 0.46. This is due to the higher attack ratio (31%), which improved separability and detection. Thus, this result suggests that tree-based ensemble models demonstrated stronger capability in identifying unseen vehicular anomalies. However, Naive Bayes and LR had the weakest overall results among the models. As seen in Table 2, the average F1-scores for Naive Bayes (0.18) and LR (0.16) were the lowest compared to the tree-based algorithms RF (0.24), XGBoost (0.26) and Extra Trees (0.25). Additionally, Table 2 presents the statistical comparison between benchmark and zero-day average F1-scores using Welch's t-test. The results indicate a statistically significant drop in F1-score performance between benchmark and zero-day attacks across RF, XGBoost, LR, and Extra Trees. However, Naive Bayes showed no statistical significance due to its consistently poor performance. Lastly, for several unseen attacks, ROC-AUC values remained close to 0.5, indicating difficulty in separating benign and unseen attack distributions. Although some unseen attacks achieved high ROC-AUC values, low F1-scores indicate poor threshold-level classification due to severe class imbalance and excessive false-positive predictions.

Moreover, Figure 2 shows the heatmap of the models' F1-scores across the unseen attacks. The DoS attack seems to be the easiest to classify, as all models achieved comparable results. On the other hand, data replay and reversed heading were considered difficult attacks to classify across all models, with all models achieving approximately 0.18 F1-scores, while LR achieved 0.12. However, LR achieved its strongest zero-day performance on the sudden-stop attack, with an F1-score of 0.26. As for traffic congestion sybil, RF, XGBoost, and Extra Trees clearly outperformed both Naive Bayes and LR, with XGBoost achieving an F1-score of 0.47 compared to 0.08 for Naive Bayes and LR.

4. Discussion

Comparing benchmark performance with zero-day evaluation, it is apparent that the models performed better on known attacks than on unseen attacks.

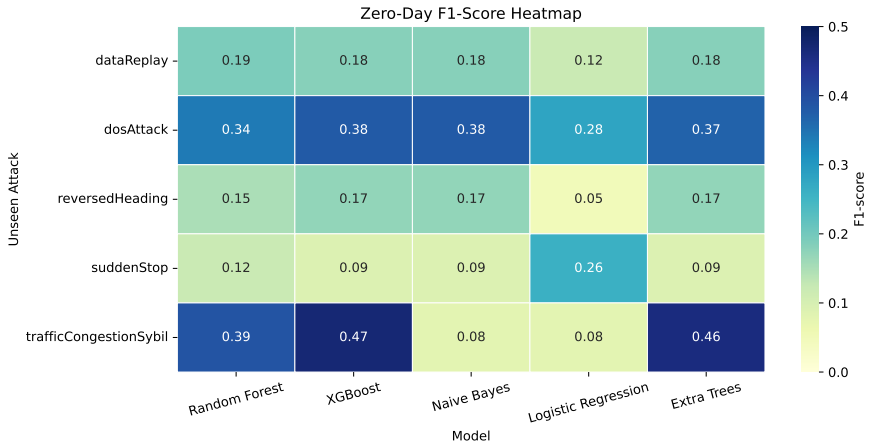

FIGURE 2. Zero-day F1-score heatmap across unseen attacks.

TABLE 2. Comparison between benchmark and zero-day average F1-scores using Welch's t-test

Model	Benchmark F1	Zero-Day Avg F1	p-value	Significant
Random Forest	0.65	0.24	0.0015	Yes
XGBoost	0.70	0.26	0.0035	Yes
Naive Bayes	0.15	0.18	0.6078	No
Logistic Regression	0.33	0.16	0.0218	Yes
Extra Trees	0.69	0.25	0.0032	Yes

TABLE 3. Performance Comparison of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Random Forest	0.83 ± 0.0011	0.59 ± 0.0027	0.71 ± 0.0020	0.65 ± 0.0012	0.89 ± 0.0007
XGBoost	0.86 ± 0.0011	0.65 ± 0.0025	0.77 ± 0.0032	0.70 ± 0.0023	0.92 ± 0.0011
Naive Bayes	0.75 ± 0.0006	0.28 ± 0.0015	0.11 ± 0.0017	0.15 ± 0.0018	0.56 ± 0.0005
Logistic Regression	0.56 ± 0.0010	0.25 ± 0.0003	0.52 ± 0.0014	0.33 ± 0.0003	0.56 ± 0.0009
ExtraTrees	0.87 ± 0.0012	0.69 ± 0.0034	0.68 ± 0.0029	0.69 ± 0.0027	0.89 ± 0.0007

Hence, this shows that the models struggle to generalise to unseen attacks. Moreover, model accuracy alone is insufficient to evaluate zero-day attack detection performance [12]. In imbalanced datasets, accuracy can be misleading because a model may achieve high accuracy despite poor attack-detection capability. Therefore, the F1-score provides a more reliable measure of model performance in such classification problems [3]. In Table 3 and 4, there is a significant drop in F1-score across all models, which illustrates the poor generalisation of the models to unseen attacks. Moreover, the results demonstrate high recall but low precision. This indicates that the models are misclassifying numerous benign samples as malicious while detecting a high number of attacks, leading to an increased false positive rate. Although detecting these unseen attacks is crucial for safety, excessive false alarms undermine the system's reliability and practicality [2].

Additionally, comparing the individual models, tree-based models were the strongest in classifying unseen attacks. Models such as RF, XGBoost and Extra

TABLE 4. Selected Leave-One-Attack-Out Zero-Day Results

Model	Unseen Attack	Accuracy	Precision	Recall	F1-score	ROC-AUC	Benign% / Attack%
Random Forest	dataReplay	0.46 ± 0.0265	0.11 ± 0.0008	0.62 ± 0.0324	0.19 ± 0.0006	0.55 ± 0.0018	90% / 10%
	dosAttack	0.44 ± 0.0127	0.23 ± 0.0051	0.61 ± 0.0182	0.34 ± 0.0071	0.51 ± 0.0083	77% / 23%
	reversedHeading	0.50 ± 0.0108	0.09 ± 0.0038	0.46 ± 0.0161	0.15 ± 0.0061	0.48 ± 0.0135	91% / 9%
	suddenStop	0.35 ± 0.0319	0.06 ± 0.0029	1.00 ± 0.0000	0.12 ± 0.0052	1.00 ± 0.0002	96% / 4%
	trafficCongestionSybil	0.52 ± 0.0081	0.32 ± 0.0047	0.50 ± 0.0049	0.39 ± 0.0029	0.53 ± 0.0035	69% / 31%
XGBoost	dataReplay	0.12 ± 0.0011	0.10 ± 0.0004	0.98 ± 0.0038	0.18 ± 0.0008	0.52 ± 0.0034	90% / 10%
	dosAttack	0.24 ± 0.0028	0.23 ± 0.0005	0.98 ± 0.0050	0.38 ± 0.0008	0.47 ± 0.0054	77% / 23%
	reversedHeading	0.12 ± 0.0047	0.09 ± 0.0011	0.96 ± 0.0123	0.17 ± 0.0019	0.47 ± 0.0091	91% / 9%
	suddenStop	0.06 ± 0.0043	0.05 ± 0.0002	1.00 ± 0.0000	0.09 ± 0.0004	0.96 ± 0.0053	96% / 4%
	trafficCongestionSybil	0.32 ± 0.0045	0.31 ± 0.0015	0.97 ± 0.0228	0.47 ± 0.0043	0.53 ± 0.0077	69% / 31%
Naive Bayes	dataReplay	0.16 ± 0.0082	0.10 ± 0.0002	0.94 ± 0.0084	0.18 ± 0.0002	0.51 ± 0.0011	90% / 10%
	dosAttack	0.27 ± 0.0004	0.24 ± 0.0002	0.94 ± 0.0012	0.38 ± 0.0003	0.50 ± 0.0007	77% / 23%
	reversedHeading	0.16 ± 0.0100	0.09 ± 0.0001	0.91 ± 0.0124	0.17 ± 0.0003	0.49 ± 0.0010	91% / 9%
	suddenStop	0.09 ± 0.0025	0.05 ± 0.0001	0.96 ± 0.0050	0.09 ± 0.0002	0.56 ± 0.0022	96% / 4%
	trafficCongestionSybil	0.68 ± 0.0033	0.35 ± 0.0107	0.05 ± 0.0091	0.08 ± 0.0145	0.53 ± 0.0073	69% / 31%
Logistic Regression	dataReplay	0.79 ± 0.0140	0.10 ± 0.0003	0.14 ± 0.0186	0.12 ± 0.0065	0.50 ± 0.0008	90% / 10%
	dosAttack	0.61 ± 0.0190	0.25 ± 0.0016	0.33 ± 0.0361	0.28 ± 0.0119	0.51 ± 0.0008	77% / 23%
	reversedHeading	0.84 ± 0.0068	0.06 ± 0.0016	0.04 ± 0.0058	0.05 ± 0.0044	0.45 ± 0.0005	91% / 9%
	suddenStop	0.92 ± 0.0053	0.22 ± 0.0096	0.32 ± 0.0235	0.26 ± 0.0050	0.83 ± 0.0045	96% / 4%
	trafficCongestionSybil	0.67 ± 0.0041	0.28 ± 0.0027	0.05 ± 0.0091	0.08 ± 0.0131	0.48 ± 0.0005	69% / 31%
Extra Trees	dataReplay	0.18 ± 0.0063	0.10 ± 0.0005	0.91 ± 0.0043	0.18 ± 0.0008	0.53 ± 0.0017	90% / 10%
	dosAttack	0.27 ± 0.0052	0.23 ± 0.0008	0.93 ± 0.0065	0.37 ± 0.0010	0.50 ± 0.0041	77% / 23%
	reversedHeading	0.18 ± 0.0036	0.09 ± 0.0008	0.88 ± 0.0092	0.17 ± 0.0015	0.48 ± 0.0113	91% / 9%
	suddenStop	0.07 ± 0.0012	0.05 ± 0.0001	1.00 ± 0.0000	0.09 ± 0.0001	1.00 ± 0.0003	96% / 4%
	trafficCongestionSybil	0.35 ± 0.0044	0.31 ± 0.0019	0.87 ± 0.0058	0.46 ± 0.0027	0.50 ± 0.0046	69% / 31%

Trees had the most competitive F1-scores [9, 8]. This suggests that ensemble-based tree models are more effective at capturing complex nonlinear relationships within vehicular communication data. Moreover, it is important to avoid overfitting of the models. For instance, RF achieved strong benchmark results but weak performance against zero-day attacks. This indicates that RF learned attack-specific patterns rather than general malicious behaviours. It should be noted that Naive Bayes and LR had the lowest average zero-day F1-score performance. Compared to [16], their results confirm that Naive Bayes underperforms in such detection systems. This could be due to the complexity and correlation among the features used to identify malicious attacks, as Naive Bayes assumes independence among features, whereas LR assumes linear separation between groups. Naive Bayes assumes conditional independence among input features [14], such that

$$P(x_1, x_2, \ldots, x_d|y) = \prod_{j=1}^{d} P(x_j|y).$$

However, vehicular features such as position, speed, acceleration and heading are naturally correlated. This implies that $\mathrm{Cov}(x_i, x_j) \neq 0$ for several feature pairs, violating the independence assumption and helping explain the weak F1-score performance of Naive Bayes compared with tree-based ensemble models. As a result, both models struggled to generalise to unseen attacks. Moreover, based on the ROC-AUC results, most models had scores around 0.5, indicating they could not reliably distinguish between benign and unseen attacks.

Traditional IDS studies often focus on improving model performance in detecting malicious cyber-attacks, which may not accurately reflect real-world deployment conditions. In practice, attackers continuously introduce evolving attack strategies that the IDS may have never been trained on before. Therefore, the findings highlight the necessity of zero-day evaluation frameworks. LOAO provides a controlled framework for evaluating IDS generalisation to unseen attacks.

4.1. **Limitations and future work.** Although tree-based ensemble models demonstrated comparatively stronger performance, overall zero-day detection capability remains limited due to poor generalisation and high false-positive rates. Future work should explore other ML models, such as Support Vector Machines (SVMs) and advanced deep learning architectures, in zero-day attack scenarios to improve robustness against evolving vehicular cyber-attacks and reduce false-positive detections. Although this study was limited to the VeReMi NextGen dataset, the dataset provides diverse attack behaviours suitable for controlled zero-day IDS evaluation. Future work should validate the proposed framework across additional VANET datasets and real-world vehicular environments [13, 10].

5. Conclusion

This study investigated the generalisation capability of multiple ML models for zero-day attack detection in VANET environments using a Leave-One-Attack-Out evaluation strategy. These findings demonstrated that benchmark results may be misleading in identifying the robustness of IDS in real-world applications [17, 4].

Although tree-based models such as RF, XGBoost and Extra Trees achieved strong benchmark results, their performance declined during LOAO evaluation. Moreover, Naive Bayes and LR performed poorly compared to the other models in both scenarios. These findings highlight the limitations of traditional ML models in handling unseen vehicular cyber-attacks and emphasise the need for more robust and adaptive IDS strategies. Overall, the results demonstrate that high benchmark accuracy alone does not guarantee reliable zero-day detection performance in VANET environments.

References

1. Fuad A. Ghaleb, Faisal Saeed, Mohammad Al-Sarem, Bander Ali Saleh Al-rimy, Wadii Boulila, AEM Eljialy, Khalid Aloufi, and Mamoun Alazab, *Misbehavior-aware on-demand collaborative intrusion detection system using distributed ensemble learning for vanet*, Electronics **9** (2020), no. 9, 1411, https://doi.org/10.3390/electronics9091411.
2. Mohammed A Abdelmaguid, Hossam S Hassanein, and Mohammad Zulkernine, *Samm: Situation awareness with machine learning for misbehavior detection in vanet*, Proceedings of the 17th international conference on availability, reliability and security, 2022, https://doi.org/10.1145/3538969.3543788, pp. 1–10.
3. Asaad Balla, Mohamed Hadi Habaebi, Elfatih AA Elsheikh, Md Rafiqul Islam, and FM Suliman, *The effect of dataset imbalance on the performance of scada intrusion detection systems*, Sensors **23** (2023), no. 2, 758, https://doi.org/10.3390/s23020758.
4. Ryma Douas and Soumia Kharfouchi, *Wavelet neural network modeling under censoring*, Gulf Journal of Mathematics **23** (2026), no. 1.
5. Mohamed Amine Ferrag, Leandros Maglaras, Sotiris Moschoyiannis, and Helge Janicke, *Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study*, Journal of Information Security and Applications **50** (2020), 102419, https://doi.org/10.1016/j.jisa.2019.102419.
6. Artur Hermann, Jan-Niklas Remmers, Dennis Eisermann, Benjamin Erb, and Frank Kargl, *Veremi nextgen: A dataset for evaluating misbehavior detection systems in vanets*, 2026, https://doi.org/10.5281/zenodo.19665762.

7. A. Jalili, M. Gheisari, J. A. Alzubi, C. Fernández-Campusano, F. Kamalov, and S. Moussa, *A novel model for efficient cluster head selection in mobile WSNs using residual energy and neural networks*, Measurement: Sensors **33** (2024), 101144.
8. Firuz Kamalov, *Sensitivity analysis for feature selection*, 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, December 2018, pp. 1466–1470.
9. Firuz Kamalov, David Santandreu Calonge, Linda Smail, Dilshod Azizov, Dimple R. Thadani, Theresa Kwong, and Amara Atif, *Evolution of AI in education: Agentic workflows*, arXiv preprint arXiv:2504.20082 (2025).
10. Firuz Kamalov, Ho-Hon Leung, and Aswani Kumar Cherukuri, *Keep it simple: Random oversampling for imbalanced data*, 2023 Advances in Science and Engineering Technology International Conferences (ASET) (Dubai, United Arab Emirates), IEEE, February 2023, pp. 1–4.
11. Anthony Kelly and Marc Anthony Johnson, *Investigating the statistical assumptions of naïve bayes classifiers*, 2021 55th annual conference on information sciences and systems (CISS), IEEE, 2021, https://doi.org/10.1109/CISS50987.2021.9400215, pp. 1–6.
12. Dominik Kus, Eric Wagner, Jan Pennekamp, Konrad Wolsing, Ina Berenice Fink, Markus Dahlmanns, Klaus Wehrle, and Martin Henze, *A false sense of security? revisiting the state of machine learning-based industrial intrusion detection*, Proceedings of the 8th ACM on Cyber-Physical System Security Workshop, 2022, https://doi.org/10.1145/3494107.3522773, pp. 73–84.
13. Amril Nazir, Rohan Mitra, Hana Sulieman, and Firuz Kamalov, *Suspicious behavior detection with temporal feature extraction and time-series classification for shoplifting crime prevention*, Sensors **23** (2023), no. 13, 5811.
14. Bharameeporn Phatcharathada and Patchanok Srisuradetchai, *Randomized feature and bootstrapped naive bayes classification*, Applied System Innovation **8** (2025), no. 4, 94, https://doi.org/10.3390/asi8040094.
15. Aparup Roy, Debotosh Bhattacharjee, and Ondrej Krejcar, *Improving internet of vehicles research: A systematic preprocessing framework for the veremi dataset*, Data in Brief **60** (2025), 111599, https://doi.org/10.1016/j.dib.2025.111599.
16. Aekta Sharma and Arunita Jaekel, *Machine learning based misbehaviour detection in vanet using consecutive bsm approach*, IEEE Open Journal of Vehicular Technology **3** (2021), 1–14, https://doi.org/10.1109/OJVT.2021.3138354.
17. Zio Souleymane, Bernard Lamien, Mohamed Beidari, and Tougri Inoussa, *Numerical simulation of pollutants transport in groundwater using deep neural networks informed by physics*, Gulf Journal of Mathematics **16** (2024), no. 2, 337–352.

Department of Electrical Engineering, Canadian University Dubai, Dubai, UAE.

Email address: 20230003036@students.cud.ac.ae