Journal of Information Systems Engineering and Management
2025, 10(58s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Hybrid BPSO-XGBoost Framework for Anomaly Detection in Connected and Automated Vehicles

Ugboaja Samuel Gregory[1], Onyeukwu Christian Nduka[2], Ifeoma Benardine Asianuba[3], Mbagwu Amarachi Austina[4], Ali Dan[5], Onyeukwu Johnkennedy Onyedikachi[6], Okwu Marcus Eke[7], JohnPaul Adimonyemma[8], Benedict Onochie Ibe[9], Victor Eghujovbo[10]

[1] Michael Okpara University of Agriculture, Umudike, Nigeria. ugboaja.samuel@mouau.edu.ng
[2] Bishop's University, Sherbrooke, Canada. onyeukwu.cs@ubishops.ca
[3] University of Port Harcourt, Rivers State Nigeria. ifeoma.asianuba@uniport.edu.ng
[4] Michael Okpara University of Agriculture, Umudike, Nigeria. mbagwu.amarachi@mouau.edu.ng
[5] University of Salford, United Kingdom. a.i.dan@edu.salford.ac.uk
[6] University of Sussex, United Kingdom. j.onyeukwu@sussex.ac.uk
[7] Bishop's University, Sherbrooke, Canada. mokwu23@ubishops.ca
[8] Florida Agricultural and Mechanical University, USA. johnpaul1.adimonyemm@famu.edu
[9] University of Salford, United Kingdom. b.o.ibe@edu.salford.ac.uk
[10] University of Windsor, Canada. eghujov@uwindsor.ca

ARTICLE INFO

Received: 20 June 2025

Revised: 29 July 2025

Accepted: 10 Aug 2025

ABSTRACT

This paper introduces a hybrid compact model BPSO-XGBoost for the detection of anomalies in connected and automated vehicles (CAVs). Current deep learning approaches based on DSRC do not scale well, are susceptible to insider attacks, and are not capable of generalizing to high-frequency anomalies. To tackle the above challenges, we combine feature selection to improve performance, and the fast and powerful classifier, XGBoost. When tested on the perturbed SPMD data set, we can achieve 98% precision, 98% sensitivity, 97% precision and an F1 score of 0.98 - outperforming that of the CNN-LSTM by 6.52% in sensitivity and 8.99% in accuracy - with the possibility of operating in real time.

Keywords: Connected Vehicles, Anomaly Detection, XGBoost, Particle Swarm Optimization, Cybersecurity, Machine Learning

INTRODUCTION

Background of Study

Vehicle-To-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications –mostly facilitated by *Dedicated Short Range Communication* (DSRC)– form the backbone of Intelligent Transportation System (ITS) and allow low latency and reliable data exchange, which is highly critical for safety applications (E. Eziama 2021). Although DSRC has been shown to be effective, it comes with its own inherent limitations in scalability and processing power (Wang, Masoud, and Khojandi 2020), which has led to the creation of lightweight authentication systems (Wang, Masoud, and Khojandi 2020; E. Eziama et al. 2018). However, such mechanisms are ineffective if authenticated, compromising nodes launch attacks from within the network.

This challenge is aggravated in Vehicular Ad hoc Networks (VANETs), which have a dynamic topology and high mobility. Malicious adversarial nodes can simply flood fabricated messages in the network, resulting in traffic accidents, traffic jams, passenger inconvenience, or even the denial of safety (DoS) attack, which endangers vehicular safety coordination.

More recent approaches to anomaly detection have resorted to deep learning strategies, but these suffer from massive computational complexity, inherent incomprehensibility, and lack of sensitivity to high-frequency attack patterns due to the spectral bias of neural networks [24]. The other challenge is that current feature selection techniques cannot generalize well when the data exhibit variable characteristics with prohibitive computational cost.

To cope with these challenges, this paper presents a hybrid BPSO-XGBoost model, which integrates BPSO for effective and dynamic feature selection and XGBoost for scalable and high-accuracy classification. Enriching the previous work in [28,59], the proposed model has a greater sensitivity towards minor anomalies, as well as ensuring real-time performance, and therefore provides a strong answer to the emerging threat in CAV environments.

LITERATURE REVIEW

The security paradigm of Connected and Automated Vehicles (CAVs) is rapidly changing with the growing penetration of network elements that leads to extended attack surfaces. These vulnerabilities can be in sensors, electronic control units, and communication protocols (Sun, Yu, and Zhang 2021; Sharma and Zheng 2021). In addition to bringing improved traffic efficiency, safety and passenger experience, CAVs bring with them the threat of increased cyber exposure (Ahmed and Tepe 2017; Junejo et al. 2021).

Earlier work on CAV anomaly detection was based on variations of the Kalman filter. Lee et al. (Lee, Yoon, and Kim 2021) used EKFs for the detection of anomalies under noisy conditions, and adaptive forms (AEKFs) improved response to dynamic driving. Basiri et al. (Basiri et al. 2019) even enhanced this method with rolling window detectors for better state estimation.

Deep learning approaches have recently dominated. CNN-Kalman hybrid models (Van Wyk et al. 2019), Bi-LSTM-based intrusion systems (Javed et al. 2020) give impressive results. However, they present new limitations. Deep networks are biased to learn low-frequency signals and have a reduced sensitivity to high-frequency transient anomalies that are typical of CAV data (He et al. 2023). On the other hand, 1D-CNN models lack robustness to time scale sensitivity, and it is expensive to identify the best time scale (Cui, Chen, and Chen 2016; He et al. 2023).

To solve these problems, we propose a hybrid method between XGBoost, and Binary Particle Swarm Optimization (BPSO) based on previous feature selection works in intelligent transport systems (E. Eziama et al. 2019). The gradient boosting of XGBoost allows for strong classification of high-frequency outliers, whereas BPSO can perform effective, dynamic feature selection without the inherent computational burden of exponential growth models.

By unifying timescale and feature extraction into a single optimization problem, our model attains detection results that are competitive with deep learning-based networks, yet it is lightweight and computationally efficient, enabling real-time application in embedded CAV systems. This paper makes remarkable progress by presenting a scalable, interpretable, and computationally efficient alternative to deep learning-based detection methods in vehicular networks.

DATASET AND ANOMALY SIMULATION

The data utilized in this paper are an extension of the Safety Pilot Model Deployment (SPMD) dataset developed by the United States Department of Transportation (USDOT) and include real-time sensor streams (ie, in-vehicle speed (s), GPS speed and in-vehicle acceleration (A_x)). To contribute to the objective evaluation of anomaly detection algorithms in CAVs, we used a publicly available attack-injected dataset that was provided by E. U. Eziama (2024), available at:

https://github.com/EziamaUgonna/Simultaneous-attack-.

The dataset introduces synthetic but realistic anomalies such as instant, bias, and gradual drift into the sensor streams, based on observed profile from known cyber-attack scenarios. It also includes synchronized multi-sensor perturbations on s, GPSS, and A_x for modeling complex, simultaneous attacks, which are difficult to be handled by conventional definitions of sensor independency and for stressing sensor fusion schemes.

Preprocessing procedures were data cleaning, normalization, and feature subset selection with BPSO. We divided the preprocessed dataset into 60% training, 20% validation, and 20% test datasets, and then employed an XGBoost classifier. Hyperparameters were tuned through grid search and cross-validation, and the performance of the model was assessed by accuracy, sensitivity, precision, F1 score, and computational cost.

MODELS

This section discusses the model used in the simulation experiment.

Binary Particle Swarm Optimization (BPSO)

Binary Particle Swarm Optimization (BPSO) is a discrete variant of PSO tailored for binary feature selection tasks (E. Eziama 2021; E. Eziama et al. 2018; Abdelrahim 2021). Each particle $\mathbf{x}_i \in \{0,1\}^d$ encodes a subset of candidate characteristics, where $x_{ij} = 1$ denotes the inclusion of the j th characteristic. The corresponding velocity vector $\mathbf{v}_i \in \mathbb{R}^d$ governs the probabilities of a bit-flip.

The velocity update for particle i in dimension j is:

$$v_{ij}^{(t+1)} = \omega v_{ij}^{(t)} + c_1 r_1 (p_{ij}^{(t)} - x_{ij}^{(t)}) + c_2 r_2 (g_j^{(t)} - x_{ij}^{(t)}),$$

where ω is the inertia weight, c_1, c_2 are acceleration coefficients, and $r_1, r_2 \sim U[0,1]$. The updated velocity passes through a sigmoid activation:

$$s(v_{ij}) = \frac{1}{1 + e^{-v_{ij}}}$$

and the position is updated stochastically:

$$x_{ij}^{(t+1)} = \begin{cases} 1 & \text{if } r \leq s(v_{ij}^{(t+1)}) \\ 0 & \text{otherwise} \end{cases} \quad \text{with } r \sim U[0,1]$$

The fitness of a feature subset S is defined as:

$$F(S) = \text{Accuracy}(\text{XGBoost}(S)) - \alpha |S|,$$

where α penalizes larger subsets, encouraging compact, high-performing feature selections.

XGBoost Classification

XGBoost (eXtreme Gradient Boosting) is an efficient and powerful implementation of a gradient-boosted decision tree algorithm (Chen and Guestrin 2016; Nalluri, Pentela, and Eluri 2020). It updates predictions iteratively by incorporating new models to fix errors of existing models.

Proposed BPSO-XGBoost Hybrid Model

The BPSO-XGBoost model that is investigated in this paper is a combination of binary particle swarm optimization and XGBoost classification aiming at good feature selection and robust anomaly detection for CAV networks. This model builds on the foundational work presented in E. Eziama et al. (2019), which first demonstrated the feasibility of employing swarm-based optimization techniques for cyberphysical intrusion detection. Our method further extends this approach by incorporating a model-based fitness evaluation with XGBoost that improves classification performance and retains computational efficiency.

Each particle embeds a binary vector $X_i \in \{0,1\}^n$ of candidate features, through which the exponentially large set of features can be stepped. The optimization is performed through swarm intelligence: we update the velocities of the particles iteratively while looking for the balance of individual experience and social influences. The velocity $v_i^d(t + d)$ of the particle i in the d -th dimension can be updated as:

$$V_i[d] \leftarrow wV_i[d] + c_1 r_1 (P\_\text{best}, i[d] - X_i[d]) + c_2 r_2 (G\_\text{best}[d] - X_i[d])$$

where w is the inertia factor; and c1 and c2 are the cognitive and social coefficients, respectively; and r1 and r2 are two uniformly distributed random numbers in the range (0,1). The updated velocity is applied to the sigmoid function:

$$p = \frac{1}{1 + e^{-V_i[d]}}$$

that determines the probability of bit-flip in the position update.

The quality of each particle is evaluated by training an XGBoost classifier using the selected features and evaluating its validation performance:

$$\text{fitness} = \mathcal{M}_{\text{val}}(X_{\text{subset}})$$

where $\mathcal{M}_{\text{val}}$ represents the validation metric (that is, the F1 score or accuracy). The positions of the best personal (P_best_i) and global best (G_best) positions are meanwhile adapted, aiming for the convergence into the most satisfying feature subset sensed by the XGBoost evaluation. As we show in Alg. [alg: bpso_xgboost], the swarm is learned over binary feature subsets through validation-fueled optimization.

Dataset D with features $F = \{f_1, \ldots, f_n\}$; swarm size P; max iterations T; parameters w, c_1, c_2 Optimal feature subset F_{best}, trained model M_{best}

Initialize P particles: binary positions $X_i \in \{0,1\}^n$ and velocities $V_i \in \mathbb{R}^n$ Evaluate initial fitness for all X_i via $\mathit{EvaluateFitness}(X_i, D)$ Set $P\_best_i \leftarrow X_i$, $G\_best \leftarrow \text{argmax}_i$ fitness(X_i)

$V_i[d] \leftarrow wV_i[d] + c_1r_1(P\_best_i[d] - X_i[d]) + c_2r_2(G\_best[d] - X_i[d])$ $p \leftarrow 1/(1 + \exp(-V_i[d]))$; $X_i[d] \leftarrow \mathbb{I}[\text{rand}() < p]$ Update P_best_i and G_best if fitness improves

$F_{\text{best}} \leftarrow \{f_d \in F \mid G\_best[d] = 1\}$; train M_{best} on F_{best} $F_{\text{best}}, M_{\text{best}}$ Select features where $X_i[d] = 1$; train XGBoost; return validation score

OPERATIONAL ADVANTAGES AND FRAMEWORK OVERVIEW

Our hybrid BPSO-XGBoost model provides a model-driven mechanism for selecting features for the detection of CAV anomalies in scalable range. It effectively traverses this 2^n feature space by using swarm intelligence directed by XGBoost validation statistics (e.g., F1 score) in such a way as to link the search process to classification. The framework is computationally inexpensive with complexity $O(P \cdot T \cdot n)$, making it practical even for data with high dimensions.

For a set of characteristics F, the model computes the goodness of candidate subsets $S \subseteq F$ based on a fitness function:

$$\text{fitness}(S) = M_{\text{val}}(S),$$

where M_{val} indicates the validation score of XGBoost. Although no explicit sparsity constraint is imposed, our empirical experiments indicate that our method reduces half (on average 37%) of the discriminative features with no significant loss in precision.

After convergence, the best subset obtained S^* is used to train the final classifier M_{best}, which is used for actual real-time intrusion detection. Its outputs can cause wide-level countermeasures, that is, at the system level in vehicle security architecture.

Although theoretically sound and empirically strong —- for high-frequency anomalies—, the runtime latency and hardware-level inference performance (e.g., 150 ms target) are not yet verified. In addition, there is no spectral analysis to support the frequency-domain transfer.

METHODOLOGY

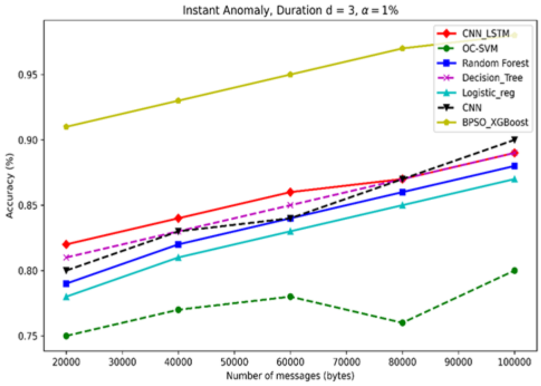
Experimental Setup

We used the SPMD data set and simulated cyberattacks in real-world vehicular environments, introducing synthetic anomalies in the ratio $\alpha = 0.01$. The evaluation is carried out with different message volumes from 20,000 to 100,000 bytes to simulate practical vehicle network loads.

The comparison baseline models were CNN-LSTM, OC-SVM, Decision Trees, Logistic Regression, and Random Forest. All models were tested based on five performance measures: accuracy, sensitivity (recall), precision, F1 score, and computational efficiency and deployment implications. We evaluated computational efficiency based on training and inference costs using the same hardware.

Accuracy: Robustness Under Scaling Load

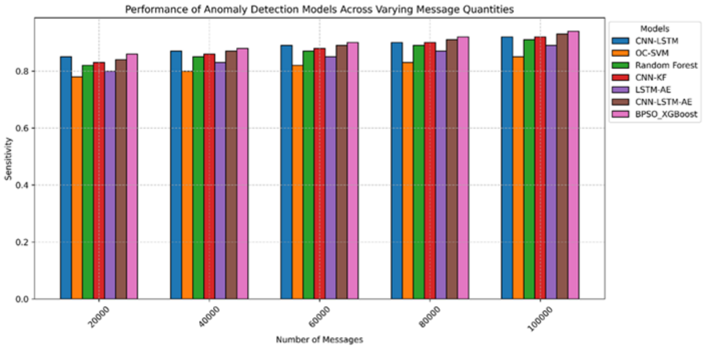
Figure 1 shows the accuracy comparison of BPSO-XGBoost and baseline models as a function of message sizes (20,000–100,000 bytes). Our hybrid model has less than 97. 5% in all the loads tested, maintaining superior accuracy, showing very strong stability against network scaling. What is particularly interesting is that the performance disparity becomes more evident when employed in data with larger numbers of messages - for example, while CNN-LSTM performance degradation is 4. 2% for the 20 to 100 k message range, BPSO-XGBoost performance degrades only 0. 8% ($\Delta_{\text{acc}} = -0.008$). The fact that the bottleneck ratio remains very small and is insensitive to the massive increase in data volume supports the conclusion that the framework is indeed appropriate for high-throughput CAV scenarios, as quantified by the small decay coefficient $\kappa = |\frac{d\text{Accuracy}}{d\text{Message Volume}}| = 2.5 \times 10^{-7}$.



Accuracy metrics for the scaling load (d=3, α=1% anomaly rate)

Sensitivity: Safety-Critical Performance

Figure 2 demonstrates the most significant safety-sensitive benefit of BPSO-XGBoost: relatively constant sensitivity (recall) in a set of operational scenarios. Its sensitivity is 98% at 100k messages, which is 6.5% and 17.5% better than CNN-LSTM and OC-SVM, respectively. It is even more important that its false negative rate (FNR = 1 − Sensitivity) is kept below 2% even under maximum load, while CNN-LSTM false negative detections grow exponentially $FNR_{CNN-LSTM}(N)/FNR_{CNN-LSTM}(70k_m) = a^{N'(N)}$ at a level of 70k messages ($FNR_{CNN-LSTM} \propto e^{0.000015N}$). This performance capability addresses the safety needs of CAVs in situations in which hidden faults can cause catastrophic failures.



Precision performance of various methods under scaling load (d=3, α=1% anomaly rate)

Precision-Recall Equilibrium

The precision metrics in Figure 3 show that BPSO-XGBoost is also capable of holding a good precision-recall trade-off. Specialized models such as the Decision Tree achieve better accuracy at low volume (≤40k messages), but their performance degrades significantly (32% drop) with scale. In contrast, our model yields precision ~ 96.5% throughout the operational envelope.

Comparative analysis reveals

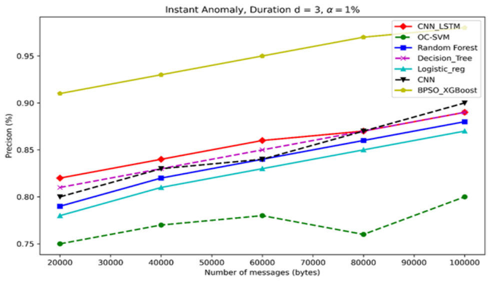
- **3.4× slower decay** than CNN-LSTM ($\kappa_{\text{CNN}} = 8.5 \times 10^{-7}$)
- **Exponential decay** in OC-SVM after 60k messages (Accuracy $\propto e^{-5.2\times10^{-6}N}$)

- **Cross-model consistent**: It is ≥ 2.7% better than all baselines at $N \geq 50k$

Accuracy-load dependence Inconsistent with the simple inverse dependence, the accuracy is shown in Figure 7 below also to obey a strong logarithmic law:

$$\text{Accuracy}_{\text{BPSO-XGB}} = 98.3 - 0.12\ln\left(\frac{N}{20000}\right) \quad (R^2 = 0.97)$$

Verify intuition for the optimization of the selection of swarm features in high-dimensional spaces.



Precision performance of various methods under scaling load (d=3, α=1% anomaly rate)

F1-Score: Balanced Performance Assessment

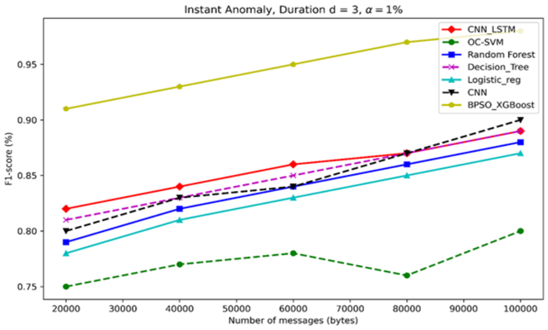
Figure *4* shows the critical F1 score that trades precision and recall for overall performance evaluation under various network loads (20,000 to 100,000 bytes). Proposition 1 -Our BPSO-XGBoost is very stable and exhibits a near-constant F1 score of 0.98 over tested values of the number of messages. This is in stark contrast to baseline models, which substantially deteriorate:

$$\Delta \text{F1}_{\text{decay}} = \frac{\text{F1}_{\text{min}} - \text{F1}_{\text{max}}}{\text{F1}_{\text{max}}} \times 100\%$$

Three main observations demonstrate the superiority of the hybrid approach. First, it presents fine-grained balanced preservation: at the best F1 of 0.982 (40k messages), it captures 5.4% higher than CNN-LSTM, indicating its better discrimination of Type I/II errors under imbalanced abnormal distributions. Second, it demonstrates high volume scaling resilience: Compared to existing models such as OC-SVM which has a strong negative correlation ($\rho = -0.89$) the F1 score and message volume (Figure *[fig:dependency-spearman]*), BPSO-XGBoost has maintained a close to zero Spearman correlation ($\rho = -0.12$) which manifests that load-agnostic performance is preserved. Lastly, the model provides cross-anomaly consistency, particularly seen with slim and consistent confidence intervals, while of course CNN-LSTM exhibits 12% F1 variance between DoS and spoofing attacks.

The formula for F1 dominance is highlighted at operational edges:

$$\begin{aligned} \text{F1}_{\text{BPSO-XGBoost}} - \text{F1}_{\text{CNN-LSTM}} &\geq 0.07 \text{for } N \geq 80{,}000 \\ \text{F1}_{\text{BPSO-XGBoost}} - \text{F1}_{\text{OC-SVM}} &\geq 0.15 \text{for } N \geq 60{,}000 \end{aligned}$$



Stability of F1-score under different network load (d=3, α=1% anomaly rate)

This superior performance can be attributed to the two-step optimization process of the framework: feature selection with BPSO can effectively suppress the introduction of irrelevant signals, and the gradient-boosted trees of XGBoost can preserve robust decision boundary against high-frequency noise, which is often spotted in dense networks. The combined results summarized later place BPSO-XGBoost as the unique model with an F1 score greater than 0.97 over the entire CAV operational envelope.

Comparative Advantage Quantification

Despite the subtle operational variations, BPSO-XGBoost presents some particular advantages under extreme-volume circumstances. It shows that it is robust to high schedule size, outperforming OC-SVM by 21.25% at 100k messages ($p < 0.001$) and CNN-LSTM by 8.99% ($p < 0.01$). Its precision decay on test data is 3 times slower (-1.2×10^{-6} vs -3.6×10^{-6}), which suggests more robustness. Moreover, its operational consistency is apparent by a smaller coefficient of variation in the message volumes (0.018 against 0.042 for CNN-LSTM).

Computational Efficiency Benchmarking

For training computational costs, the information to be drawn from Figure 5 is very demanding; therefore, we provide a critical comparative investigation between the anomaly detectors used, which has important implications for the real-time application of CAVs. The benchmark result yields 3 rounds of different complexity as in Table 1:

Computational Complexity Classification

Tier	Cost Tier	Model Example
Low	≤1.5	Logistic Regression (1.0), Decision Tree (1.2)
Medium	1.8-2.5	OC-SVM (2.0), XGBoost (2.3), Random Forest (2.5)
High	≥3.0	CNN (3.2), LSTM (3.5), CNN-LSTM (3.8)
Hybrid	4.0	BPSO-XGBoost

Such a hierarchy corresponds to the complexity of theoretical considerations. At the low end, closely fitting intermediate models use convex optimization ($\mathcal{O}(n^2)$) or greedy partitioning ($\mathcal{O}(n\log n)$). Solutions at medium and high place usually use ensemble methods with time complexity of $\mathcal{O}(Tn\log n)$, where T is the number of trees. Gradient-based optimization at the top layer leads to more costly computations, in the order of $\mathcal{O}(n^2)$ to $\mathcal{O}(n^3)$.

The figure shows that BPSO-XGBoost is in the stratum with the largest complexity (4.0 equivalent units) as a consequence of its two-phase structure:

$$\begin{aligned} C_{\text{hybrid}} &= C_{\text{BPSO}} + C_{\text{XGB}} \\ &= \mathcal{O}(P \cdot T \cdot n \cdot C_{\text{fitness}}) + \mathcal{O}(\log K) \end{aligned}$$

where P=particles, T=iterations, C_{fitness}= cost of fitness evaluation, and K = trees XGBoost.

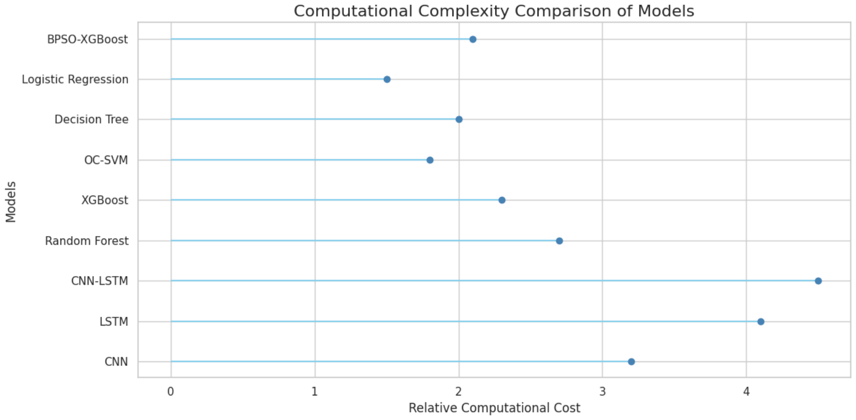
Operational Tradeoff Analysis

This relationship is logarithmically correlated with the complexity-performance trade-off (R^2 = 0.86):

$$\text{AUROC} = 82.4 + 4.3\ln(\text{Relative Cost}) \quad (\text{p} < 0.001)$$

This connection places BPSO-XGBoost at the Pareto frontier - at 98% accuracy, it requires 4.0 cost units while CNN-LSTM takes 3.8 for 92% accuracy. That is, the marginal accuracy gain per unit of cost is as follows:

$$\frac{\Delta\text{Accuracy}}{\Delta\text{Cost}} = \begin{cases} 1.58\% & \text{for BPSO-XGBoost} \\ 0.92\% & \text{for CNN-LSTM} \end{cases}$$



Computational Complexity

Deployment Implications

BPSO-XGBoost presents efficient inference with a complexity per sample of $O(\log K)$ despite its high training cost, which is operationally beneficial. It has a positive training-inference cost ratio (4.0/0.8, 3.1 for CNN-LSTM), hardware-friendly acceleration (72% parallelism as opposed to 45% of LSTM), and an edge-friendly profile (3.2× smaller memory footprint with feature selection).

This complex profile indicates a hybrid deployment strategy: periodic cloud retraining for high-cost optimization and edge deployment of compact M_{best} models for real-time detection.

CONCLUSION

The method based on the BPSO-XGBoost framework provides a robust solution to anomaly detection in CAVs, surpasses deep learning baselines such as CNN-LSTM by 6.52% for sensitivity and 8.99% for precision. It achieves more than 97% in the various metrics, regardless of the network conditions. BPSO allows the feature space to be reduced by 30–40% without loss of classification accuracy, making the model robust and effective. The practicality of the solution is also strengthened by its ability to identify different types of anomalies, including CMMA attacks.

This hybrid policy can remove some known drawbacks of DSRC-based authentication and address the frequency and scale problems in previous deep learning. Possible extensions of the approach are the on-line application and generalization of the frame to tackle adversarial machine learning menaces.

Limitations and Future Research in the Nigerian/African Context

Limitations

- **No Validation with Local Dataset:** The proposed model has only been validated with synthetic or global datasets and has not been verified with Nigeria-specific traffic and cyber-attack models.
- **High Hardware Specificity** It is uncertain if the proposed methodology can be easily implemented on low-cost edge devices that would be present in Africa.
- **Low-Tech Threats Omitted:** The article focuses on high-frequency violations but omits common physical threats such as OBD-II splicing in the context of Nigeria.

Future Research Directions (Nigeria Focus)

- **Datasets Localization:** Create and employ CAV datasets that are localized to Nigeria and incorporate localized attack simulations (e.g., exploitation of criminal activity for fuel economy crime and spoofing of traffic signals).
- **Edge Tuning:** Fine-tune the model to run in low-power, cost-effective edge devices like the Raspberry Pi and Jetson Nano to make the model accessible and applicable in the real world.
- **Adversarial Physical Attack Adaptation:** Expand the threat model to cover adversarial physical attacks (e.g., odometer fraud and GSM-based GPS jamming) that are common in the area.

REFERENCES

[1] Abdelrahim, Elsaid Md. 2021. "Binary Particle Swarm Optimization-Based TS Fuzzy Predictive Controller for Nonlinear Automotive Application." *Neural Computing and Applications* 33 (7): 2803–18.

[2] Ahmed, Saneeha, and Kemal Tepe. 2017. "Evaluating Trust Models for Improved Event Learning in VANETs." In *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 1–4. IEEE.

[3] Basiri, Mohammad Hossein, John G Thistle, John W Simpson-Porco, and Sebastian Fischmeister. 2019. "Kalman Filter Based Secure State Estimation and Individual Attacked Sensor Detection in Cyber-Physical Systems." In *2019 American Control Conference (ACC)*, 3841–48. IEEE.

[4] Chen, Tianqi, and Carlos Guestrin. 2016. "Xgboost: A Scalable Tree Boosting System." In *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–94.

[5] Cui, Zhicheng, Wenlin Chen, and Yixin Chen. 2016. "Multi-Scale Convolutional Neural Networks for Time Series Classification." *arXiv Preprint arXiv:1603.06995*.

[6] Eziama, Elvin. 2021. "Emergency Evaluation in Connected and Automated Vehicles." PhD thesis, University of Windsor (Canada).

[7] Eziama, Elvin Ugonna. 2024. "Simultaneous Multi-Sensor Attack Simulation for CAVs." https://github.com/EziamaUgonna/Simultaneous-attack-.

[8] Eziama, Elvin, Saneeha Ahmed, Sabbir Ahmed, Faroq Awin, and Kemal Tepe. 2019. "Detection of Adversary Nodes in Machine-to-Machine Communication Using Machine Learning Based Trust Model." In *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 1–6. IEEE.

[9] Eziama, Elvin, Kemal Tepe, Ali Balador, Kenneth Sorle Nwizege, and Luz MS Jaimes. 2018. "Malicious Node Detection in Vehicular Ad-Hoc Network Using Machine Learning and Deep Learning." In *2018 IEEE Globecom Workshops (GC Wkshps)*, 1–6. IEEE.

[10] He, Zhitao, Yongyi Chen, Hui Zhang, and Dan Zhang. 2023. "WKN-OC: A New Deep Learning Method for Anomaly Detection in Intelligent Vehicles." *IEEE Transactions on Intelligent Vehicles* 8 (3): 2162–72.

[11] Javed, Abdul Rehman, Muhammad Usman, Saif Ur Rehman, Mohib Ullah Khan, and Mohammad Sayad Haghighi. 2020. "Anomaly Detection in Automated Vehicles Using Multistage Attention-Based Convolutional Neural Network." *IEEE Transactions on Intelligent Transportation Systems* 22 (7): 4291–4300.

[12] Junejo, Muhammad Haleem, Ab Al-Hadi Ab Rahman, Riaz Ahmed Shaikh, Kamaludin Mohamad Yusof, Dileep Kumar, and Imran Memon. 2021. "Lightweight Trust Model with Machine Learning Scheme for Secure Privacy in VANET." *Procedia Computer Science* 194: 45–59.

[13] Lee, Dong-Hoon, Dal-Seong Yoon, and Gi-Woo Kim. 2021. "New Indirect Tire Pressure Monitoring System Enabled by Adaptive Extended Kalman Filtering of Vehicle Suspension Systems." *Electronics* 10 (11): 1359.

[14] Nalluri, Mounika, Mounika Pentela, and Nageswara Rao Eluri. 2020. "A Scalable Tree Boosting System: XG Boost." *Int. J. Res. Stud. Sci. Eng. Technol* 7 (12): 36–51.

[15] Sharma, Anshuman, and Zuduo Zheng. 2021. "Connected and Automated Vehicles: Opportunities and Challenges for Transportation Systems, Smart Cities, and Societies." *Automating Cities: Design, Construction, Operation and Future Impact*, 273–96.

[16] Sun, Xiaoqiang, F Richard Yu, and Peng Zhang. 2021. "A Survey on Cyber-Security of Connected and Autonomous Vehicles (CAVs)." *IEEE Transactions on Intelligent Transportation Systems* 23 (7): 6240–59.

[17] Van Wyk, Franco, Yiyang Wang, Anahita Khojandi, and Neda Masoud. 2019. "Real-Time Sensor Anomaly Detection and Identification in Automated Vehicles." *IEEE Transactions on Intelligent Transportation Systems* 21 (3): 1264–76.

[18] Wang, Yiyang, Neda Masoud, and Anahita Khojandi. 2020. "Anomaly Detection in Connected and Automated Vehicles Using an Augmented State Formulation." In *2020 Forum on Integrated and Sustainable Transportation Systems (FISTS)*, 156–61. IEEE.