Cureus Journal of Computer Science

Open Access Original Article

Enhanced Anomaly Detection in Wireless 5G Networks With Hybrid Learning Technique Using AWID3 Dataset

Nasim Dashtifard [1] , Haitham Mahmoud [1] , Moad Idrissi [2] , Nouh Elmitwally [3]

1. Networks, Birmingham City University, Birmingham, GBR 2. School of Computing and Data Science, Oryx Universal College | Liverpool John Moores University, Doha, QAT 3. Computing, Birmingham City University, Birmingham, GBR

Corresponding authors: Nasim Dashtifard, nasim.dashtifard@mail.bcu.ac.uk, Haitham Mahmoud, haithamhassanmahmoud@gmail.com

Received 04/23/2025
Review began 04/24/2025
Review ended 06/27/2025
Published 06/30/2025

DOI: https://doi.org/10.7759/s44389-025-05486-0

Abstract

In recent years, the expansion of the Internet of Things and 5G networks has significantly increased wireless traffic, heightening the risk of cyberattacks. Intrusion detection systems have become essential for safeguarding wireless networks by providing real-time threat detection and response. This study presents a comprehensive review and implementation of machine learning-based techniques for detecting various types of wireless attacks, with a focus on improving detection accuracy through ensemble learning. The AWID3 dataset, based on the IEEE 802.11 standard, was used for experimentation. The study was conducted in multiple phases: (1) evaluating six machine learning algorithms (random forest, J48, naïve Bayes, logistic regression, decision tree, and deep neural networks) using three feature selection methods (information gain, gain ratio, and chi-squared); (2) developing a hybrid ensemble model by integrating the strengths of deep neural network, random forest, XGBoost, and LightGBM, with logistic regression as a meta-classifier; and (3) validating performance using key metrics: accuracy, precision, recall, and F1-score. The proposed hybrid model achieved a peak accuracy of 99.75%, outperforming benchmark models in the literature. These results demonstrate the superior performance and robustness of the proposed hybrid approach. By addressing multiple network layers and leveraging ensemble learning, this research highlights the critical role of hybrid models in achieving reliable and accurate intrusion detection for wireless environments.

Categories: Cyber-Physical Systems, Data Analysis, Machine Learning (ML)
Keywords: cybersecurity, machine learning, 5g networks, feature selection, hybrid model

Introduction

The rapid growth in wireless devices demands a unified and scalable platform to manage increasing traffic. To address this, fifth-generation (5G) communication technology has emerged, offering advanced capabilities that meet the needs of next-generation wireless networks. 5G enables the integration of concepts such as the Internet of Things (IoT), revolutionising connectivity across platforms, software, people, and devices [1]. It supports advanced radio access technologies, including non-orthogonal multiple access, mm-Wave, and massive multiple-input multiple-output, enhancing communication between user equipment like IoT devices. However, traditional security mechanisms such as authentication, encryption, and firewalls struggle to balance efficiency and protection in these dynamic environments [2,3].

Due to the growing attack surface in wireless environments, intrusion detection systems (IDSs) have become critical. Several studies have used AWID2 [4-6] and AWID3 datasets to detect wireless network attacks. These datasets serve as benchmarks for identifying malicious traffic using machine learning techniques.

Islam and Allayear [3] applied a K-nearest neighbour classifier to AWID3, targeting address resolution protocol, de-authentication, AMOK, and authentication request attacks. Their results showed higher accuracy in detecting address resolution protocol attacks and recommended expanding feature sets for improved recall, particularly by including the wlan_ra (MAC address) feature. Moreover, Chatzoglou et al. [7] compared AWID2 and AWID3, finding that AWID3 yielded better results for legacy flooding attacks. With only 16 features, their model achieved up to 99.55% accuracy in shallow learning and 97.55% in deep learning for seven specific attacks, including Deauth, Kr00k, and Evil Twin. Furthermore, Saini et al. [5] addressed limitations in AWID3's WPA2 focus by generating their own dataset using a testbed. They developed a hybrid IDS system that combined signature-based and machine learning-based detection, achieving 99% accuracy while planning to expand support for WPA3. Furthermore, Chatzoglou et al. [8] examined 802.11 and non-802.11 application-layer attacks using AWID3. They applied decision trees, LightGBM, and Bagging, reaching 99% accuracy with feature set conflation.

How to cite this article
Dashtifard N, Mahmoud H, Idrissi M, et al. (June 30, 2025) Enhanced Anomaly Detection in Wireless 5G Networks With Hybrid Learning Technique Using AWID3 Dataset. Cureus J Comput Sci 2 : es44389-025-05486-0. DOI https://doi.org/10.7759/s44389-025-05486-0

Da Silva et al. [6] focused on evil twin attacks, applying data balancing techniques like RUS and K-means SMOTE. Using six machine learning algorithms and K-fold cross-validation, they showed that LightGBM achieved a low false positive rate of 0.00602 when using all columns. The "wlan.fc.protected" column was found to have the most influence on the detection based on the feature understanding.
Furthermore, Sethuraman et al. [9] developed an intrusion detection method using kernel density estimation and hidden Markov model for detecting Evil Twin, Wi-Fi phishing, and injection attacks. Their system achieved 98% accuracy and required no extra hardware or protocol changes, demonstrating the ability to detect unknown attacks effectively. Furthermore, Salah and Elsoud [10] performed a multi-phase evaluation on AWID3, addressing nominal, numeric, and binary classes across five attack types. Using logistic regression, they achieved up to 99% accuracy, with the lowest results being 89.1%, 60%, and 86.7% across different stages. Furthermore, Kasongo and Sun [11] proposed an IDS approach integrating feed-forward deep neural network (DNN) with a Wrapper Feature Extraction Unit. They applied their model to both UNSW-NB15 and AWID datasets, achieving up to 99.77% accuracy for multiclass classification on AWID, validating the model's adaptability and detection performance. In addition, Yang et al. [12] explored intrusion detection in IoT traffic using active learning. Their human-in-the-loop approach prioritised informative data points and significantly improved detection performance compared to supervised learning, although they noted that its application in wireless IoT remains at an early stage.

These works highlight the growing relevance of intelligent IDS, particularly those leveraging machine learning for intrusion detection in wireless environments. Machine learning enables early identification of anomalies, including zero-day attacks, through continuous learning and adaptive detection [3,4,12]. Techniques like anomaly detection and feature selection further enhance model accuracy and efficiency. Hybrid models, or ensemble methods, integrate strengths from multiple algorithms to boost anomaly detection. This study proposes such a hybrid model to achieve high classification accuracy by evaluating different machine learning algorithms and feature selection techniques to design a robust security solution for wireless networks.

Materials And Methods

This section outlines the process for the framework for intrusion detection in wireless networks using hybrid methods in machine learning techniques, as shown in Figure *1*. The process includes data pre-processing, data split, both detection techniques, and evaluation.

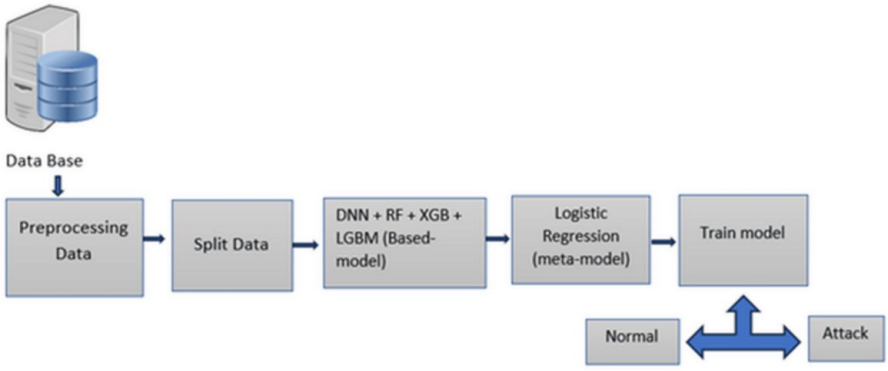

FIGURE 1: Overview of the proposed hybrid methodology for anomaly detection, illustrating a two-layered detection process: first, individual base models (DNN, Random Forest, XGBoost, LightGBM) analyse input features to detect potential anomalies; second, a meta-classifier (Logistic Regression) integrates their outputs to produce a final, more accurate decision

DNN, Deep Neural Network; XGBoost, eXtreme Gradient Boosting; LightGBM, Light Gradient Boosting Machine

Dataset

This study works on the AWID3 dataset. The AWID project provides real-world wireless attacks to provide a solution towards identifying a resilient security approach for wireless network traffic data. Although some modifications and corrective actions have taken place, the vulnerabilities in the last version have been neglected. Since the security in wireless networks has long remained unsolved. Therefore, outside

security approaches need to be considered as essential components in the 802.11 standard for cellular networks for preventing attacks [6]. The AWID dataset was introduced in 2016, and the latest version was released in 2021 called AWID3. The AWID dataset replicates 802.11 standard network attacks, including the WPA and WPA2 security protocols [3,10]. WPA3 security protocol attacks have been considered in this latest version of the dataset. It involves various traffic data from cellular network traffic captured in a controlled laboratory environment. This dataset has 254 features, including 253 generic features and one additional feature for labelling. This provides an overall of 36,913,503 instances of both (6,526,404) malicious and (30,387,099) normal traffic. It contains 13 different attack types from all the OSI layers model, including MAC address layer attacks to application layer attacks. In this study, all CSV files of different attacks are merged into one file, and after preprocessing, 31 features are considered for exploration.

Attacks

The AWID3 dataset extends the original AWID dataset by including attacks from higher-layer protocols as well as newly emerging threats. According to Čermák et al. [13], the attacks in AWID3 are categorised into four groups:

- 802.11-Specific Attacks: These are MAC layer (Layer 2) attacks that exploit vulnerabilities in the IEEE 802.11 protocol to disrupt wireless communication [14,15]. They typically involve repeated request flooding or manipulation of control frames. This category includes well-known attacks such as Deauthentication, Disassociation, Reassociation, Rogue Access Point (AP), KRACK, and Kr00k, all of which are comprehensively represented in the AWID3 dataset.

- Local Node Attacks: Originating from benign nodes within the local network, these attacks target system resources at higher layers, particularly the application layer. Examples include SSH Brute Force, Botnet, and Malware [16,17].

- External Node Attacks: These attacks are launched from outside the local network, targeting local clients by exploiting vulnerabilities in Internet-connected systems. Included in this category are SSDP Amplification and SQL Injection attacks [18].

- Multi-Layer Attacks: These complex attacks span multiple layers, typically involving at least two layers and unknown network architectures such as the Internet [19,20]. AWID3 includes Evil Twin and Website Spoofing under this category. Table *1* provides a detailed breakdown of each attack type included in the dataset.

Attack	Normal traffic	Malicious traffic
Deauth	1,587,527	38,942
Disas	1,938,585	75,131
(Re)Assoc	1,838,430	5,502
Rogue AP	1,971,875	1,310
Krack	1,388,498	49,990
Kr00k	2,708,637	186,173
SSH	2,428,688	11,882
Botnet	3,169,167	56,891
Malware	2,181,148	131,611
SQL Injection	2,595,727	2,629
SSDP	2,641,517	5,456,395
Evil Twin	3,673,854	104,827
Website spoofing	2,263,446	405,121
Total	30,387,099	6,526,404

TABLE 1: Table of attacks

SSDP, Simple Service Discovery Protocol; SSH, Secure Shell

Preprocessing

The AWID3 dataset, which consists of WPA2-Enterprise and Protected Management Frames, includes 13 distinct types of wireless network attacks ranging from legacy deauthentication to more advanced threats such as Kr00k and malware, as shown in Table *2*. To manage the large volume of data, which includes over 70 CSV files per attack type, the preprocessing phase was crucial. First, the individual CSV files were merged for each attack type into a consolidated dataset. This large dataset was further split based on attack types to optimise resource usage. For data cleaning, Python 3.9 was used to handle missing values, perform data normalisation, and encode categorical variables. Missing values were addressed using a combination of strategies: filling them with the mean, median, or mode and in cases where this was not applicable, removing rows with missing data. This process ensured the dataset was clean and ready for analysis. To facilitate machine learning model training, categorical features, such as attack labels, were converted to numeric values using label encoding. Binary features were transformed into 0 or 1 for consistency. Additionally, non-essential columns, such as timestamps and addresses, were removed to reduce dimensionality and minimise noise in the data.

Name of traffic	Convert to Numeric
Normal	0
(Re)Assoc	1
Botnet	2
Deauth	3
Disas	4
Evil Twin	5
Kr00k	6
Malware	7
Krack	8
Rogue AP	9
SQL Injection	10
SSDP	11
SSH	12

TABLE 2: Convert nominal to numeric

SSDP, Simple Service Discovery Protocol; SSH, Secure Shell

Finally, the dataset was normalised using MinMaxScaler to ensure uniformity in feature ranges, which improved model performance. The data was then split into training (65%) and testing (35%) sets based on the distribution of attack types, ensuring balanced representation across all classes. This preprocessing pipeline provided a clean, structured dataset optimised for machine learning algorithms.

Feature selection

Feature selection, also known as attribute selection, is a process of extracting the most relevant features from the dataset [10] and then applying machine learning algorithms for the better performance of the model. Three distinct feature selection techniques have been employed: Information Gain, Gain Ratio, and Chi-Square. They are selected based on the most relevant features of the AWID3 dataset. Each method represents one aspect of selecting features based on their contribution to predicting the pattern and improving the model's performance. The dataset train file is loaded into a Pandas DataFrame. Although there is no direct function for Information Gain in scikit-learn, we calculate it manually or use libraries like sklearn combined with custom implementations.

Hybrid model

In this study, a hybrid machine learning model is developed using a stacking ensemble approach to maximise detection accuracy. The hybrid architecture integrates several complementary models, leveraging their strengths while compensating for their limitations. Specifically, the DNN component captures complex nonlinear relationships within the data. Random Forest (RF), configured with 200 estimators, provides robust and stable predictions through bagging. Meanwhile, the gradient boosting models, XGBoost and LightGBM, contribute high accuracy and efficient learning capabilities, making them well suited for handling imbalanced and high-dimensional datasets. Collectively, these models form a unified ensemble that enhances overall predictive performance.

The hybrid model consists of four different base models, including DNN, RF, XGBoost and LightGBM, to gain the capabilities of each of them in the ensemble. A DNN is a type of artificial neural network that can learn complex patterns in large datasets with the help of multiple layers of neurons. This model is suitable for high-dimensionality datasets and requires modelling complex dependencies. In this hybrid model, DNN is implemented using the MLPClassifier (Multi-Layer Perceptron) from the scikit-learn library. The architecture of the hidden layer is (100, 50) neurons. Selecting two hidden layers with 100 and 50 neurons enabled the model to analyse the large number of patterns and interact effectively with the data. A wide range of patterns can be provided by the first layer with 100 neurons and enabling the model to learn more abstract features by purifying those selected patterns, provided by the second layer with 50 neurons. This architecture provides nonlinear interactions between features in the data, and DNN can capture and

model intricate relationships. In addition, it provides an effective balance between model complexity and computational efficiency. Selecting too many neurons may lead to overfitting, especially with limited data.

RF is a type of decision tree able to handle both categorical and continuous variables. It is implemented by an RF Classifier from scikit-learn. This model builds multiple decision trees during training and makes a class with classes of the individual trees. This hybrid model configures an RF with 200 estimators (trees). Each tree has a random subset of features that cause enhancement and decrease the risk of overfitting. The reason for selecting a 200 estimator is to keep the balance between model performance and computational cost. This number is often chosen by testing. Increasing the number of trees generally improves the model's performance by reducing variance. However, after a certain point, the computational cost increases and the performance gain goes down significantly. It shows that adding more trees beyond this number does not significantly improve accuracy but does increase training time and resource usage. With 200 trees, the RF model can also provide reliable feature selection, leading to selecting features which are most influential in making predictions.

Gradient boosting XGBoost was configured by 500 Estimators, Learning Rate 0.01, and Max Depth 6. Each estimator corrects the errors of the earlier ones. As a result, more estimators generally lead to better performance. However, using more than this number may cause overfitting. By using a high number of estimators and a low learning rate (0.01), the model can learn gradually, and it will take significant time to estimate the pattern, which often results in better generalisation of unseen data. However, the learning rate increases the learning time, and it should be specified wisely since the model has sufficient time to build strong decision rules. This is particularly effective when combined with a larger number of estimators. The maximum depth of each tree is set to 6 to prevent the model from becoming too complex and overfitting the training data. A max depth of 6 is often chosen because it allows the model to analyse important interactions between features without becoming complex.

LightGBM is generally used for large datasets because the histogram-based algorithm of LightGBM does not need heavy computation and gains high accuracy with a high learning rate in deeper trees. In LightGBM, the Estimator is configured 500, the learning rate is 0.01 with a max depth of 6 for LightGBM mirrors that of XGBoost to leverage its efficiency advantages while maintaining similar predictive capabilities. Although LightGBM and XGBoost are both gradient-boosting models, LightGBM is optimised for speed and efficiency aspects. In this method, both XGBoost and LightGBM have the same configuration with different implementations. This strategy helps us to use the power of both gradient boosting methods, which enhances the ensemble's overall performance. The hybrid model consists of four different base models, including DNN, RF, XGBoost, and LightGBM, to gain the capabilities of each of them in the ensemble.

The meta-algorithm in this method is the Logistic Regression model, which serves as the meta-model (final_estimator). Logistic Regression is a linear model, which means that it is predicted based on the weighted sum of input features. The prediction in the hybrid method comes from base models. The predictions of all based models are combined and sent to Logistic Regression as a meta-method to make a final prediction. One of the advantages of the meta-model context is simplicity. This characteristic makes the hybrid model less complex, which could lead to overfitting. The coefficients of Logistic Regression are straightforward to interpret. Each coefficient shows the level of the importance of the contribution of each model in the prediction of the result. Each based model gets the weight for contributing to the hybrid model, which helps us understand the ensemble's behaviour.

Results

Experimental setup

The experiments were conducted on a dataset containing 13 distinct attack types to evaluate the performance of multiple machine learning algorithms in a multi-class classification setting. The selected models include RF, J48, Decision Trees, Naïve Bayes, Logistic Regression, and DNN. Each algorithm was tested using three feature selection techniques, Information Gain, Gain Ratio, and Chi-Squared, to assess their impact on model performance. The evaluation was based on key metrics: accuracy, precision, recall, and F1-score. All experiments were performed under consistent conditions to ensure fair comparison across models and feature selection methods.

Traditional models

We present results from experiments on a dataset with 13 distinct attack types, as in Table 3, aiming to assess the effectiveness of various machine learning algorithms with different feature selection methods. The models tested include RF, J48, Decision Trees, Naïve Bayes, Logistic Regression, and DNN, each evaluated using Information Gain, Gain Ratio, and Chi-Squared techniques. Performance was measured using accuracy, precision, recall, and F1-score, key metrics for evaluating multi-class classification models.

Model	Feature selection approach: Info Gain				Feature selection approach: Gain Ratio				Feature selection approach: Chi-Squared			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Random Forest	0.9963	0.9963	0.9963	0.9963	0.983	0.9811	0.983	0.9802	0.9939	0.9937	0.9939	0.9938
treesJ48	0.9967	0.9975	0.997	0.997	0.9813	0.981	0.981	0.981	0.996	0.997	0.997	0.9975
Naïve Bayes	0.7095	0.9668	0.7147	0.8482	0.6681	0.9265	0.6765	0.7584	0.6865	0.9656	0.6841	0.7865
Logistic Regression	0.9576	0.9468	0.9687	0.9578	0.9734	0.9672	0.9724	0.9765	0.9548	0.9475	0.9585	0.9544
Decision Tree	0.9964	0.9402	0.944	0.942	0.9964	0.9391	0.9439	0.9415	0.9965	0.9423	0.9408	0.9414
Deep Neural Network	0.9859	0.9847	0.9994	0.9885	0.9064	0.874	0.9185	0.8985	0.9645	0.9635	0.9665	0.9678

TABLE 3: Results of our proposed traditional algorithms (of single detection approach)

Hybrid model

We evaluated six traditional machine learning models, RF, Decision Tree (treesj48), Naïve Bayes, Logistic Regression, and DNN, against the proposed hybrid model (as shown in Table *4*, Figures *2* and 3). Each model was tested using three feature selection techniques, with the best results reported. RF with Info Gain achieved a high accuracy of 0.9963. Decision Tree (treesj48) also performed well, with an accuracy and F1-score of 0.997 using the same technique. Naïve Bayes and Logistic Regression showed weaker performance, with Naïve Bayes dropping to 0.7095 in one case. Compared to all, the hybrid model consistently outperformed individual models, achieving a top accuracy of 0.9973 and strong performance across all metrics.

Table *5* presents the results of the hybrid model compared with both base models and existing literature. This hybrid combines DNN, RF, XGBoost, and LightGBM as base models, with Logistic Regression as the meta-model. Predictions from all base models are aggregated and passed to the Logistic Regression layer for the final output.

Algorithm	Feature selection approaches	Accuracy	Precision	Recall	F1-Score
Random Forest	Info Gain	0.9963	0.9963	0.9963	0.9963
treesJ48	Info Gain	0.9967	0.9971	0.9976	0.997
Naïve Bayes	Info Gain	0.7095	0.9662	0.7152	0.874
Logistic Regression	Gain Ratio	0.9734	0.9685	0.9765	0.9775
Decision Tree	Chi-Squared	0.9965	0.9423	0.9408	0.9414
Deep Neural Network	Info Gain	0.9859	0.9827	0.9975	0.9882
Proposed Hybrid Model	None	0.9973	0.9972	0.9973	0.9972

TABLE 4: Comparing and evaluating hybrid results with traditional algorithms

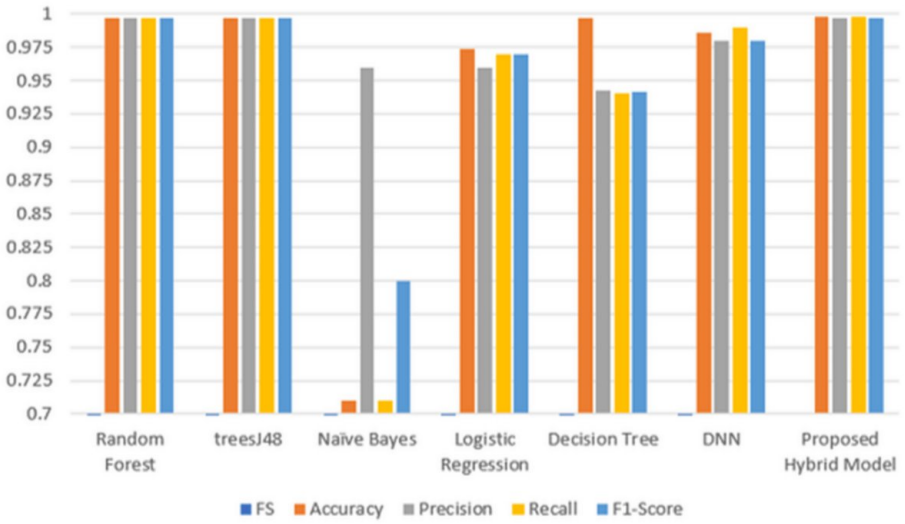

FIGURE 2: Comparison of hybrid with traditional models, where x-axis is the used machine learning and the y-axis refers to % of different metrics

DNN, Deep Neural Network; FS, Feature Selection

Benchmarking

Algorithm	Feature selection	Accuracy	Precision	Recall	F1-Score
Chatzoglou et al. [7]	LightGBM	0.9656	0.9725	0.9481	0.9599
	Bagging	0.967	0.9684	0.9503	0.9591
Chatzoglou et al. [8]	LightGBM	0.9942	0.9989	0.9922	0.9955
Da Silva et al. [6]	LightGBM	0.99259	0.98898	0.9963	0.99262
	XGBoost	0.99153	0.98741	0.99577	0.99157
	LightGBM (Optimized)	0.99286	0.98898	0.99683	0.99289
Proposed Hybrid Model	None	0.9973	0.9972	0.9973	0.9972

TABLE 5: Comparison of the proposed hybrid model with other existing studies that rely on the hybrid approaches

XGBoost, eXtreme Gradient Boosting; LightGBM, Light Gradient Boosting Machine

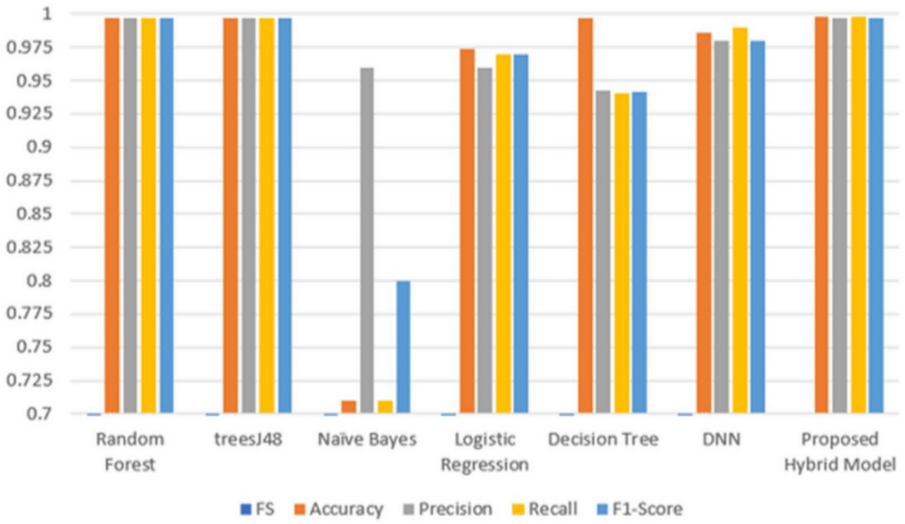

FIGURE 3: Comparison of hybrid models with other papers' hybrid models, where x-axis is the existing studies and the y-axis refers to % of different metrics

DNN, Deep Neural Network; FS, Feature Selection

Chatzoglou et al. [7] divided the dataset into feature groups and applied a hybrid model combining LightGBM and Bagging. Their best accuracy was 0.9656 using LightGBM and 0.9670 using Bagging on a combined feature set created by merging sets 16 and 17 with engineered features, which are lower than those achieved by our proposed hybrid model. Moreover, Chatzoglou et al. [8] evaluated LightGBM on the AWID2 and AWID3 datasets using four subsets. The highest accuracy obtained was 99.42% on AWID3, which is still 0.003 lower than the result achieved by our hybrid model.

Da Silva et al. [6] evaluated three hybrid models, LightGBM, XGBoost, and Optimised LightGBM, on AWID3 twin attacks. Optimised LightGBM achieved 0.9928 on the fourth column set; XGBoost and LightGBM achieved 0.9915 and 0.9925, respectively. Our hybrid model outperformed all, with an accuracy of 0.9973.

Overall, while prior studies used models like DNN, RF, and gradient boosting on AWID3 with strong results, our hybrid approach of integrating DNN, RF, XGBoost, and LightGBM achieves higher accuracy and robustness, setting a new benchmark in the literature.

Discussion

Traditional models

We evaluated six traditional machine learning models: RF, J48, Decision Tree, Naïve Bayes, Logistic Regression, and DNN using three feature selection methods: Information Gain, Gain Ratio, and Chi-Squared. The best results per model were reported. RF and J48 (Info Gain) showed strong performance with accuracies of 0.9963 and 0.9967, respectively. In contrast, Naïve Bayes showed weaker performance (accuracy as low as 0.7095). Overall, while some models performed well, variability across metrics highlighted the limitations of relying on single classifiers.

Hybrid model

To overcome these limitations, we developed a hybrid model combining DNN, RF, XGBoost, and LightGBM as base models, with Logistic Regression as a meta-classifier. This ensemble achieved superior results across all metrics: accuracy (0.9973), precision (0.9972), recall (0.9973), and F1-score (0.9972), outperforming all traditional models.

Benchmarking with literature

Compared to recent studies, the proposed model achieves the highest accuracy. For instance, Chatzoglou et al. [7] reported 0.9656 with LightGBM and 0.9670 with Bagging, while Chatzoglou et al. [8] achieved 0.9942. Da Silva et al. [6] reached up to 0.9928 with Optimised LightGBM. None surpassed our hybrid model's 0.9973 accuracy.

Conclusions

This study aims to enhance the predictive capabilities of traditional machine learning models for detecting a wide range of cyberattacks in 5G wireless networks. To this end, a hybrid ensemble model was developed, integrating DNN, RF, XGBoost, and LightGBM as base learners, with Logistic Regression serving as the meta-classifier. The model was implemented in Python and benchmarked against six traditional algorithms using three established feature selection techniques: Information Gain, Gain Ratio, and Chi-Squared.

Experimental results clearly illustrate the effectiveness of the proposed system. The hybrid model achieved an outstanding accuracy of 0.9973, surpassing the performance of all individual base models and state-of-the-art approaches in the literature. Notably, this exceeds state-of-the-art results of a peak accuracy of 0.9670 using a LightGBM-Bagging hybrid, and 99.42% using standalone LightGBM on the AWID3 dataset. It also outperforms the state-of-the-art of 0.9928 with an optimized LightGBM approach. These findings highlight the superior accuracy and robustness of our hybrid system for wireless intrusion detection.

While the current model demonstrates strong performance, several directions can further enhance its practical deployment. Future work will explore alternative meta-models (e.g., K-Nearest Neighbours or other ensemble strategies), increase base model diversity, and incorporate advanced feature selection methods. Furthermore, the model will be extended to detect previously unseen attack types, thereby improving resilience. Finally, optimizing training and inference efficiency will be crucial to enabling real-time deployment in resource-constrained 5G environments. Real-world applications include deployment in smart city surveillance networks, industrial IoT systems, 5G-enabled healthcare devices, and wireless edge environments such as autonomous transportation and remote critical infrastructure monitoring, where fast and accurate intrusion detection is essential for maintaining operational security and system availability.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Haitham Mahmoud, Nasim Dashtifard

Acquisition, analysis, or interpretation of data: Haitham Mahmoud, Nasim Dashtifard, Moad Idrissi, Nouh Elmitwally

Critical review of the manuscript for important intellectual content: Haitham Mahmoud, Moad Idrissi, Nouh Elmitwally

Supervision: Haitham Mahmoud, Nouh Elmitwally

Drafting of the manuscript: Nasim Dashtifard

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Chettri L, Bera R: A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. IEEE Internet of Things Journal. 2020, 7:16-32. 10.1109/jiot.2019.2948888
2. Brownlee J: Introduction to Dimensionality Reduction for Machine Learning. Machine Learning Mastery, Vermont, Australia; 2020.
3. Islam T, Allayear SM: Capable of classifying the tuples with wireless attacks detection using machine learning. Intelligent Computing Systems. Brito-Loeza C, Martin-Gonzalez A, Castañeda-Zeman V, Safi A (ed): Springer, Cham; 2022. 1569:1-16. 10.1007/978-3-030-98457-1_1
4. Chatzoglou E, Kambourakis G, Kolias C, Smiliotopoulos C: Pick quality over quantity: Expert feature selection and data preprocessing for 802.11 Intrusion Detection Systems. IEEE Access. 2022, 10:64761-64784.

10.1109/access.2022.3183597
5. Saini R, Halder D, Baswade AM: RIDS: Real-time intrusion detection system for WPA3 enabled enterprise networks. GLOBECOM 2022 - 2022 IEEE Global Communications Conference. 2022, 43-48. 10.1109/GLOBECOM48099.2022.10001501
6. da Silva LM, Andreghetti VM, Romero RAF, Branco KRLJC: Analysis and identification of evil twin attack through data science techniques using AWID3 dataset. MLMI '23: Proceedings of the 6th International Conference on Machine Learning and Machine Intelligence. 2023, 128-135. 10.1145/3635638.3635665
7. Chatzoglou E, Kambourakis G, Smiliotopoulos C, Kolias C: Best of both worlds: Detecting application layer attacks through 802.11 and non-802.11 features. Sensors. 2022, 22:5633. 10.3390/s22155633
8. Chatzoglou E, Kambourakis G, Kolias C: Empirical evaluation of attacks against IEEE 802.11 enterprise networks: The AWID3 dataset. IEEE Access. 2021, 9:34188-34205. 10.1109/ACCESS.2021.3061609
9. Sethuraman SC, Dhamodaran S, Vijayakumar V: Intrusion detection system for detecting wireless attacks in IEEE 802.11 networks. IET Networks. 2019, 8:219-232. 10.1049/iet-net.2018.5050
10. Salah Z, Elsoud EA: Enhancing intrusion detection in 5G and IoT environments: A comprehensive machine learning approach leveraging AWID3 dataset [PREPRINT]. Preprints. 2023, 10.20944/preprints202307.1565.v1
11. Kasongo SM, Sun Y: A deep long short-term memory based classifier for wireless intrusion detection system. ICT Express. 2020, 6:98-103. 10.1016/j.icte.2019.08.004
12. Yang K, Ren J, Zhu Y, Zhang W: Active learning for wireless IoT intrusion detection. IEEE Wireless Communications. 2018, 25:19-25. 10.1109/MWC.2017.1800079
13. Čermák M, Svorenčík Š, Lipovský R, Kubovič O: KR00K - CVE-2019-15126: Serious vulnerability deep inside your Wi-Fi encryption. ESET White Paper. 2020.
14. Kolias C, Kambourakis G, Stavrou A, Gritzalis S: Intrusion detection in 802.11 networks: Empirical evaluation of threats and a public dataset. IEEE Communications Surveys & Tutorials. 2015, 18:184-208. 10.1109/COMST.2015.2402161
15. Tahsien SM, Karimipour H, Spachos P: Machine learning based solutions for security of Internet of Things (IoT): A survey. Journal of Network and Computer Applications. 2020, 161:102630. 10.1016/j.jnca.2020.102630
16. Khraisat A, Gondal I, Vamplew P, Kamruzzaman J: Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity. 2019, 2:20. 10.1186/s42400-019-0038-7
17. Hamroun C, Fladenmuller A, Pariente M, Pujolle G: Intrusion detection in 5G and Wi-Fi networks: A survey of current methods, challenges and perspectives. IEEE Access. 2025, 13:40950-40976. 10.1109/access.2025.3546338
18. Singh NJ, Hoque N, Singh KR, Bhattacharyya DK: Botnet-based IoT network traffic analysis using deep learning. Security and Privacy. 2023, 7:e355. 10.1002/spy2.355
19. Shukla P, Krishna CR, Patil NV: Iot traffic-based DDoS attacks detection mechanisms: A comprehensive review. The Journal of Supercomputing. 2024, 80:9986-10043. 10.1007/s11227-023-05843-7
20. Rani SVJ, Ioannou II, Nagaradjane P, Christophorou C, Vassiliou V, Yarramsetti H: A novel deep hierarchical machine learning approach for identification of known and unknown multiple security attacks in a D2D communications network. IEEE Access. 2023, 11:95161-95194. 10.1109/access.2023.3308036