Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 3(3) May 2025
DOI: https://dx.doi.org/10.4314/jobasr.v3i3.5

An Optimized Random under Sampling Boosting Model for Intrusion Detection in Wireless Sensor Network

Fatima Suleiman[1], Umar Iliyasu[2], Mukhtar Abubakar[3*] and Sanusi Abdul Sule[4]
[1, 2]Department of Computer Science, Federal University Dutsin-Ma, Katsina, Nigeria
[3]Department of Software Engineering, Federal University Dutsin-Ma, Katsina, Nigeria
[4]Department of Information Technology, Federal University Dutsin-Ma, Katsina, Nigeria
*Corresponding Author Email: mabubakar2@fudutsinma.edu.ng

ABSTRACT
The rapid growth of Information Technology and the COVID-19 pandemic have significantly increased internet usage worldwide, leading to a surge in cyber threats. Network Intrusion Detection Systems (NIDS) are essential for monitoring and mitigating unauthorized intrusions, but class imbalance among attack types often hampers performance, particularly for minority samples. This research proposes an Optimized Random Under sampling Boosting Model using Particle Swarm Optimization algorithm to enhance the detection and classification of five intrusion types: DOS, Probe, R2L, U2R, and Normal. The model outperformed state-of-the-art models such as JN8, KNN, and GA–LSTM–RNA, achieving near-perfect performance with approximately 100% accuracy for all attack types. In terms of precision, the model achieved 99.990% for DOS, 99.886% for Probe, 97.689% for R2L, 73.684% for U2R, and 99.990% for Normal. The recall rates were similarly impressive, with 99.993% for DOS, 99.943% for Probe, 99.329% for R2L, 93.333% for U2R, and 99.886% for Normal. The F1-scores were 99.956% for DOS, 99.914% for Probe, 98.502% for R2L, 82.353% for U2R, and 99.938% for Normal. In general, the proposed ORUSBEM model was able to effectively detect each type of attack, including the normal type. The results revealed the effectiveness of the developed technique in enhancing the accuracy and efficiency of the intrusion detection system. The result is beneficiary in improving the security of various applications by detecting and preventing malicious attacks. The model executed 100 iterations in 185.1 seconds, indicating significant optimization improvements. While BPSO corroborate to be effective in selecting the optimal feature subset, other optimization techniques such as Genetic Algorithms (GA), Ant Colony Optimization (ACO), or Differential Evolution (DE) can be used to compare and optimize the results in the future. However, focusing on enhancing U2R detection and reducing execution time is also an area that needs serious attention. Additionally, testing the model in real-life applications is crucial for strengthening cyber security in industries and organizations.

Keywords:
Wireless,
Sensor,
Network,
Threat,
Cyber,
Under sampling

INTRODUCTION

The advancement of Information Technology has led to increase access of internet and network use across the globe. These services serve to be of paramount importance to companies and other industries to carry out their activities irrespective of distance, time and at a convenient place worldwide. The rapid increase of the said online activities on the internet has resulted to various attacks to the network from the users(Kumar et al., 2021). The aforesaid problem has narrowed the attention of various researchers to come up with an optimize models for detecting such attacks and mitigating them. Network Intrusion Detection Systems (NIDS) are security systems designed to monitor, detect and mitigate dynamically against intrusions by unauthorized users (Cao et al., 2022). This is carried out using data collected from several computer nodes which are analyzed to ascertain the threat level in the network. Several techniques were developed to achieve intrusion detection. This includes; rule-based systems,

which performance heavily reliant on security professionals' rules. The method of enciphering set of instructions is costly and inefficient due to the large volume of network traffic. In order to tackle the challenges of rule based system, data mining is employed wireless sensor network for NIDS(Ieracitano et al., 2018; Tan et al., 2019).

Wireless Sensor Networks (WSNs) have become integral components of modern technological ecosystems, facilitating applications ranging from environmental monitoring to industrial automation. However, the pervasive deployment of these networks has exposed them to a myriad of security challenges, with potential attackers exploiting vulnerabilities for various malicious purposes. One prevalent threat is the Denial of Service (DoS) attack, where adversaries overwhelm the network by flooding it with spurious traffic, thereby disrupting communication and rendering sensors ineffective. Additionally, WSNs are susceptible to node replication attacks, where adversaries clone legitimate nodes to inject false information into the network Grover et al., (2016). Ensuring the security of wireless sensor networks is imperative, requiring robust encryption, authentication mechanisms, and intrusion detection systems to safeguard sensitive data and maintain the reliability of these critical infrastructures. As the adoption of WSNs continues to grow, addressing these security challenges becomes paramount to harness the full potential of these pervasive technologies. Furthermore, Wireless sensor network is nowadays a field that has gained serious attention due to the increasing reliance in information and communication technology and limitless access to data. These increasing online activities have increased the risk of malicious attacks on users in the networks. Network traffic datasets used for analyzing threats are characterized by a large number of features and class imbalance in the attack types such as the NSL-KDD dataset which increases the complexity of an intrusion detection model and increases the error rate of minority class (Muhuri et al., 2020; Mulyanto et al., 2021). Various algorithms were implemented to detect intrusions in a network, however, the imbalance nature of attack types (such as: Denial of Service (DoS), user to root (U2R), Remote to Local (R2L) and Probing attacks) increases the error rate in identifying these attacks due to model bias that occurs during training. This is because only few algorithms considered class imbalance in their implementation but could still not improve accuracy for the detection of the minority class. Also, most of the authors do not consider the time of execution which at times can take many minutes if not hours to execute the processes. Hence, this research work proposes the development of an optimized random under sampling boosting model for intrusion detection in wireless sensor network. The research is expected to improve the detection of minority intrusion attacks while minimizing the execution time for the processes.

Various Researchers have been carried out in the field of Network Intrusion Detection Systems (NIDS), among them includes the work of (Nagaraja et al., 2020), developed a network anomaly detection related to feature transformation. The method adopted distance function for carrying feature clustering to achieve feature transformation. In order to recognize anomalous traffic, the model was subjected to detect whether its anomalous or not based on the two datasets used such as KDD (41 and 19 features) and NSL-KDD (41 features). Some popular classifiers used are; J48 and KNN. A 10-fold cross-validation performance evaluation with accuracy, precision, recall and F-score were evaluated. The findings obtained revealed an overall improvement in the accuracy of the proposed model with feature transformation for all the three datasets used. However, the efficiency of the model for attack classes like U2R and R2L with low sample sizes compared to Normal, Dos and Probe was low. For example, the F-score obtained by J48 and KNN classifiers for NSL-KDD-41 datasets for U2R 0.5682 and 0.7184 respectively. Also, the F-score obtained by J48 and KNN for R2L are 0.9460 and 0.9150. This result indicates the need for data balancing to improve the attacks with few sample sizes

(Tang et al. 2020) presented an IDS based on LightGBM and an AutoEncoder. The LightGBM, which is a boosting, based algorithm, is used for feature selection. It works by combining several decision trees to assign scores to each feature. This process leads to a selection of 21 features of the NSL-KDD which totals to 102 dimensions after one-hot encoding. Afterwards, the autoencoder uses the selected features for Intrusion detection. The Autoencoder works by reconstructing the features into higher order feature spaces. The error in reconstruction is used by the Auto-Encoder to set thresholds for detection. An accuracy of 89.82% in the binary class scenario as experiments were not made in the multi-class case.

(Sarumi et al. 2020) compared the performance of SVM and an association rule method for IDS. A filter-based feature selection method was applied before training. This filter is based on the computed mutual information between the feature and the label. The association rule finds patterns in the dataset that are used for classification based on user defined thresholds. The association rule uses breadth first bottom up, apriori method for identification of the features. It considers candidates patterns with n-items and tests each of them to see if its frequency exceeds the user defined thresholds. Experiments results in accuracies of 90.41% and 64.09% for the SVM and the apriori approach on the UNSW-NB15 dataset and

(77.17%, 67%) for the NSL-KDD datasets. However, the apriori approach is faster than the SVM.

(Muhuri et al., 2020) proposed an intrusion detection system to detect different attacks in NSL-KDD dataset. Genetic algorithm and Long Short Term Memory (LSTM) were used to detect different types of attack within a network. The result indicate that genetic algorithm outperformLSTM-RNN in both binary and multi-class classification when compare with SVM and RF classifiers. The model obtained an optimal accuracy of 96.81% and 99.91% for binary classification on 122 and 99 features respectively. The model revealed to have a low precision, recall and f1-score detection level of R2L and U2R attacks due to the data imbalance from the minority class.

(Wu et al., 2020) proposed a Semantic Re-encoding and Deep learning Model (SRDLM) intrusion detection technique. SRDLMmodel has been tested to increase the ability to detects anomaly detection and enhance the dataset generalizability. The model was able to enhance the accuracy and robustness of the dataset in detecting more traffic within a network. The model detects web character injection attack with an accuracy of 99% on NSL-KDD dataset. Therefore the model efficiency has been improved by 8% compared with the conventional ML algorithms

(Thapa et al., 2020) developed a robust anomaly-based network detection model for new attacks using ML and DL models. The performances of various ML and DL models were evaluated on CoburgDatasets (CIDDS). An ensemble models of both ML and DL was proposed and the model achieved an optimal accuracy of 99% for detecting new attacks.

(Zhou et al., 2020) proposes a model to increase the efficiency of IDS with high dimensional and unbalance network traffic with a framework based on feature selection and ensemble learning techniques. A CFS-BA heuristic algorithm was used for selecting the optimal subset based on the correlation between features. A C4.5, RF, forest Penalized Attributes (PA) were ensemble using voting probability for the classification. 10fold cross-validation method was used for evaluation over three intrusion detection datasets (NSL-KDD, AWID, and CIC-IDS2017). The result of their experiment shows the promising result with an accuracy of classification equal to 99.81%, 99.8% DR and 0.08% FAR with a subset of 10 features for NSL-KDD with a subset compose of 8 features AWID provide accuracy of 99.52% and 0.15% FAR. The model achieved the highest accuracy of 99.89% and DR of 99.9% on the subset of 13 feature of the CIC-IDS2017.

(Mulyanto et al. 2021) developed a focal loss network intrusion detection system (FL-NIDS) for imbalanced intrusion, FL-NIDS was applied using Deep Neural network (DNN) and Convolutional neural network (CNN) on three benchmark intrusion detection datasets. The proposed technique was compared with traditional DNN and CNN. The results shows that the FL-NIDS obtained an accuracy of 77% for binary classification with NSL-KDD dataset and 89% with UNSW-NB15 dataset. For multiclass classification, FL-NIDS obtained an accuracy of 66% and 78% for NSL-KDD and UNSW-NB15 datasets respectively. The proposed technique has not been applied to sequential tasks problems, and the performance is low which can still be improved.

Also the work of (Qazi et al. 2023) proposed Hybrid deep learning intrusion detecting model for detecting malicious attack within a network. The model outperforms the current intrusion detection models in detecting malicious attack with an average accuracy of 98.90%. More network traffic attack should be considered in the future and there is need to employ backbone network traffic to show the efficiency of the model.

Similarly the work of (Sivamohan and Sridhar 2023; Abubakat et al. 2025) come up with a bidirectional Long Short Term Memory based Explainable Artificial Intelligence. The developed framework achieved a higher accuracy of 97.2% and 95.3% for Honeypot and BSL-KDD dataset in detecting intrusion attack within a network. The developed framework provides better Safety and confidentiality within the industry but find it difficult to detect adversarial attacks in the network.

Amru et al. (2024) and Imrana et al. (2025) proposed and ensemble model and different machine learning to predict different types of network attack. It was found that XGBoosting algorithm was used to predict different classes of attack with an accuracy of 94%. One of the limitations of the ensemble model is that it cannot detect clone attack in the network.

MATERIALS AND METHODS

Data Preprocessing and Model Development

In this section, data preprocessing and the proposed optimized RUSBEM model with BPSO is described. The adopted RUSBEM model serves as the classification model for multi-class and imbalance dataset, while, the Binary Particle Swarm Optimization (BPSO) serves as the feature selector and optimizer. This approach is called wrapper-based approach where the machine learning model is required for evaluation. Random under sampling boosting (RUSBoost) ensemble model is classification models developed for classifying imbalance data into binary or multi-class. In RUSBoost, bunch of individual classification, models were trained sequentially using the mistakes of the previous model in the sequence. This approach propose n optimized Random Under sampling Boosting Ensemble Model (RUSBEM) for detection and classification of intrusion adopted from (Li, 2021). The

imbalance NSL-KDD dataset has undergo a pre-processing stage by converting all textual data to numerical equivalents, followed by normalization of features with large difference to avoid biased. All duplicates data were removed and cleaned for better classification. The dataset is divided into two, 70% for training and 30% for testing.

The filtered and normalized relevant features would be chosen for improved classification. This is necessary to reduce computational complexities and increase the efficiency of the proposed model. BPSOmetaheuristic algorithm was adopted for feature selection. This is because of its success and light weight nature as compared to other optimization feature selection algorithms.

Binary Particle Swarm Optimization (BPSO)

Binary Particle Swarm Optimization (BPSO) is a variation of Particle Swarm Optimization (PSO) that can be used for problems solutions. BPSO operates by utilizing a group of particles each representing a solution, within the search space. These particles possess both position and velocity which are updated based on rules involving the known positions of individual particles and the entire swarm. The objective is to discover the solution by guiding the particles towards promising regions within the search space.

Feature Selection with BPSO

The feature selection problem is a binary optimization task where each feature is either included or excluded from the model. The solution is depicted as a binary vector, with a value of 1 representing an included feature and 0 representing an excluded feature. Each solution is represented as a particle, where the length of the particle corresponds to the number of features, and the binary values (0 or 1) indicate which features are to be selected or discarded. Table 3.1 shows the particle representation of a feature in the solution space; where 1 indicate a chosen feature and 0 represent a rejected one as showed in Table 1 below.

Table 1: Particle representation

1	2	3	4	5	6	7	8	. . .	N
1	**1**	**1**	**0**	**0**	**1**	**0**	**1**	**. . .**	**0**

To assess the performance of each particle and determine its fitness, the mean absolute error (MAE) is calculated as illustrated in Equation 1

$$MAE = \frac{1}{S}\sum_{i=1}^{S}|A_i - L_i| \qquad (1)$$

where

S represents the total number of samples,

(A_i) denotes the actual class of sample (i)

(L_i) stands for the predicted label of sample (i).

Table 2 presents the BPSO parameters applied in the experiments. The length of each particle, or the problem dimension, corresponds to the total number of features. The maximum number of iterations ranges between 20 to 100, while the population size varies between 10 and 50. The minimum and maximum inertia weights are set to 0 and 2, respectively.

Table 2: BPSO Parameter

Parameter	**Value**
Particle length (N)	Equivalent to the total number of features
Particle population	Ranges between 10-50
Number of cycle	Between 20-100
Highest Inertia Weight (w_2)	Set to 2
Lowest Inertia Weight (w_1)	Set to 0
Acceleration constants (c1, c2)	Values of 1 and 2 respectively

Model Classification Categories

The classification task in this research is a multi-class classification where the RUSBEM was used to classify attacks in the following categories: denial of service (DoS), probe, user-to root (U2R), and remote-to-local (R2L) and normal. All these categories of attacks and normal have different number of samples making the data highly imbalanced.Using the RUSBoost algorithm, random sampling was applied to create ensemble of sequential weak learners from the training samples. Each learner sampled the data using the minority sample and classifies the data into the multiple classes. The wrongly classified samples move to the next weak learner as weighted data while repeating the same process until the number of learners 'k' is reached. The test samples are then tested on the final classifier to ascertain the performance of the ensembles on an unseen test data as presented in figure 1 below.

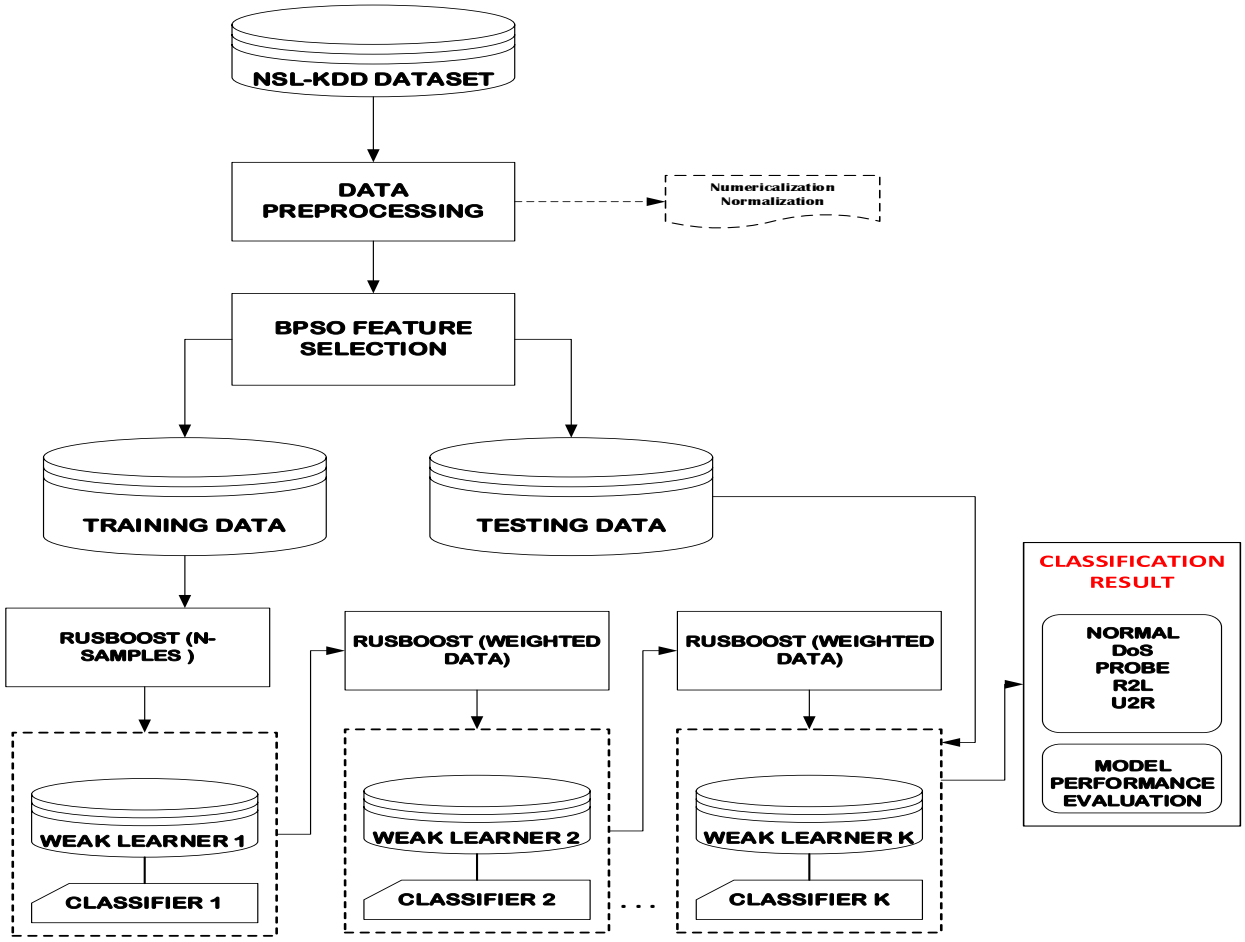

Figure 1: Proposed ORUSBEMModel

Datasets Description

The dataset is made up of five classes of attacks namely, Normal, Dos, probe, R2L and U2R attacks and has 12, 5973 as the overall number of samples. The samples comprises of normal, Dos, Probe, R2L and U2R attacks with 67343, 45927, 11656 , 995 and 52 samples respectively. The dataset comprises of 40 features and one categorical class. Table 3 below shows 4 samples with 5 features of the NSL-KDDdataset with the class showing 4 different class attributes.All the features are were preprocessed and cleaned for better classification. This will enable the removal of features with zero percentage. To do this, the number of non-zeros was calculated for each feature, and then the percentage of non-zeros was determined followed by filtering any feature with less than 30% non-zero elements. This threshold was experimental and can be reviewed during the experiment. The final features will then be normalized especially features with large difference amongst the members. Table 3 below shows the NSL-KDD dataset sample with their respective class of attack.

Table 3 NSL-KDD dataset sample

S/N	duration	srcbyte	dst_byte	land	wrg_frag	Class
1	0	491	0	0	0	'normal'
2	0	43	43	0	0	'Probe'
3	0	241	1400	0	0	'R2L'
4	0	232	8153	0	0	'Dos'

Performance Evaluation Metrics

The following metrics were used in order to assess the efficacy of the developed model as presented in equation 2 to 5 below.

Accuracy: represents the ratio of correctly predicted samples to the total number of samples.

$$\text{Accuracy} = \frac{TP+TN}{TP+FT+TN+FN} * 100 \quad (2)$$

Precision: refers to the fraction of positive predictions that are actually true positives.

$$\text{Precision} = \frac{TP}{TP+FT} * 100 \quad (3)$$

Recall: indicates the ratio of true positive cases that are correctly identified.

$$\text{Recall} = \frac{TP}{TP+FN} * 100 \quad (4)$$

F1-Score: represents the harmonic average of precision and recall

$$\text{F1 Score} = \frac{2(precision*Recall)}{Precision+Recall} * 100 \quad (5)$$

where, TP is True Positives, TN is True Negatives, FP is False Positive and FN is false negative.

RESULTS AND DISCUSSION

To test the efficacy of the developed ORUSBEM model, an experiment was performed using the optimal selected features. The ORUSBEM model was tested with 23 optimal features and 500 ensemble trees. The number of trees was chosen after some experiments with different number of trees and 500 trees produced the best results. Table 4 summarizes the model parameters, number of features selected.

Table 4: Model Investigation parameters

MODEL	Ensemble Tree Size	Feature size	%Feature Reduction
ORUSBEM	500	23	57.5%

Table 5 and figure 2 shows the performance results of each model based on accuracy, precision, recall, and F1-score for each of the attack types respectively. The accuracy score indicates the ratio of correctly predicted samples to the total number of samples. Precision indicates the fraction of positive predictions that are actually true positives.The recall score shows the proportion of accurately predicted attacks over the sum of number of real attacks. The F1-score is the harmonic mean of the precision and recall scores. The model has proved to be promising in accurately detecting different attacks categories with 100%, 99984, 99976, 99984 and 99998 for Dos, Probe, R2L and U2R and Normal attack respectively. The model obtained a precision results for detecting and classifying different types of attacks with almost 100% that is 99% and above with the exception of U2R attack with has an approximate of 74%. Moreover the Recall and FI-Socre also obtained a promising result of almost 100% for Dos, Probe, R2L and U2R with the exception of U2R which obtained 82% for FI Score.

Table 5: Model Performance

Model	Attack Types	Accuracy	Precision	Recal	F1-score	Execution Time (s)
ORUSBEM	**DOS**	1.00000	0.99990	0.99993	1.00000	185.1
	PROBE	0.99984	0.99886	0.99943	0.99914	
	R2L	0.99976	0.97689	0.99329	0.98502	
	U2R	0.99984	0.73684	0.93333	0.82353	
	NORMAL	0.99998	0.99990	0.99886	0.99938	

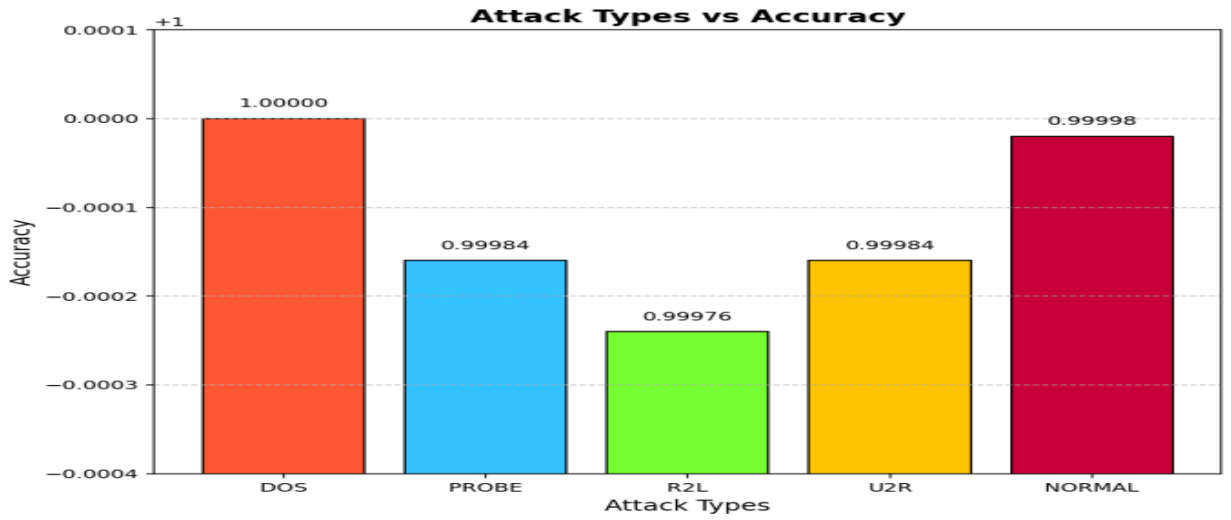

Figure 2: Attacks Type

Table 6 and figure 3 below demonstrated the performance validation results of ORUSBEM model and other research works. Among the compared methods, the proposed ORUSBEM method stands out as the most effective in terms of accuracy, precision, recall, and F1-Score for various attack types. Nagaraja et al., (2020) achieved strong performance, particularly for DOS and NORMAL attacks, using both J48 and KNN methods. Muhari et al., (2020) achieved good accuracy for some attack types but struggled with PROBE, R2L, and U2R attacks. The proposed ORUSBEM model proved to be promising in detecting and classifying different attacks compared with other models proposed by Nagaraja et al., (2020) and Muhari et al, (2020). Furthermore F1 score was employed as an additional metric to further validate the efficacy of the model. Also, time metric was utilized in the proposed ORUSBEM model to reduce time and space complexity of the computing resources which other models lacked to addressed. The Proposed model was executed in 185.1 seconds which shows the model is potent in optimizing the aforesaid results.

Table 6: Comparison between different methods

Author	Method	Attack Types	Accuracy	Precision	Recall	F1-Score	Execution time (s)
Nagaraja et al., (2020)	**J48**	**DOS**	0.99807	0.99687	0.997844		-
		PROBE	0.99588	0.98058	0.974777		-
		R2L	0.99916	0.95872	0.933668		-
		U2R	0.9997	0.69444	0.480769		-
		NORMAL	0.99598	0.99576	0.996733		-
	KNN	**DOS**	0.99768	0.99555	0.998106		-
		PROBE	0.99458	0.97705	0.964053		-
		R2L	0.99865	0.91045	0.919598		-
		U2R	0.99977	0.72549	0.711538		-
		NORMAL	0.99355	0.99377	0.994179		-
Muhari et al, (2020)	GA LSTM-RNN	**DOS**	0.9943	0.99	0.99	0.99	-
		PROBE	0.9380	0.71	0.94	0.81	-

		R2L	0.6135	0.95	0.61	0.74	-
		U2R	0.6866	0.15	0.69	0.25	-
		NORMAL	0.9947	1	0.99	1	-
Proposed	ORUSBEM	**DOS**	1.00000	1.00000	0.99993	1.00000	185.1
		PROBE	0.99984	0.99886	0.99943	0.99914	
		R2L	0.99976	0.97689	0.99329	0.98502	
		U2R	0.99984	0.73684	0.93333	0.82353	
		NORMAL	0.99934	0.9999	0.99886	0.99938	

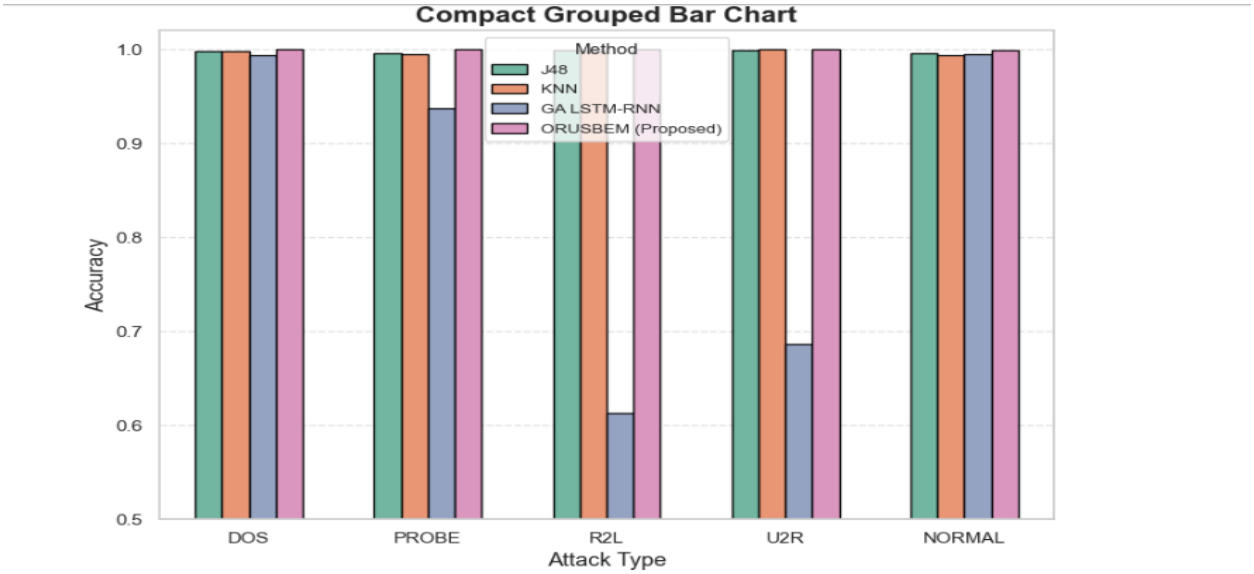

Figure 3: Comparison with SOTA

Table 7 below shows the confusion matrix of the ORUSBEM model where the confusion matrix reveals that the ORUSBEM model delivers outstanding performance across all five categories: DOS**,** PROBE**,** R2L**,** U2R**,** and NORMAL. For the DOS class, 13,778 out of 13,779 instances are accurately classified, with only 1 instance mistakenly labeled as NORMAL**.** In the PROBE category, 3,494 out of 3,499 instances are correctly identified, with minimal errors distributed across other classes. The R2L class has 296 correct classifications out of 298**,** with 2 instances misclassified as U2R. In the U2Rcategory, 14 out of 15 instances are correctly identified, with 1 misclassified as PROBE**.** Finally, the NORMAL class shows 20,180 correct predictions out of 20,203**,** with a few instances misclassified as other categories. Overall, the matrix highlight excellent performance in the major classes (DOS and NORMAL) while indicating slight challenges in rare catergories like U2R due to limited data and occasional misclassificatios.

Table 7: Confusion Matrix of the proposed ORUSBEM Model

Model	**Attack Types**	**Dos**	**Probe**	**R2L**	**U2R**	**Normal**
ORUSBEM	DOS	13778	0	0	0	1
	PROBE	1	3494	0	0	4
	R2L	0	0	296	2	0
	U2R	0	1	0	14	0
	NORMAL	10	3	7	3	20180

CONCLUSION

The development of the Optimized Random Undersampling Boost Ensemble Model (ORUSBEM) using Binary Particle Swarm Optimization (BPSO) has shown promising results in accurately detecting different types of attacks in multi-class intrusion detection. The BPSO achieved an optimal feature subset early in the search process which helped in selecting 23 features representing over 57.5% of the total features. The best features are; Duration, src_byte, land, wrg_frag, urgent, hot, logd_in, num_root, num_Fcreat, num_accF, num_Ocmd, srv_count, serror_r, srv_serror, same_srv, dst_hostSC, dst_hssr, dst_hssp, dst_hsr, and dst_hsssr.
The proposed ORUSBEM model was able to effectively detect each type of attack, including the normal type. The results revealed the effectiveness of the developed technique in enhancing the accuracy and efficiency of the intrusion detection system. The result is beneficiary in improving the security of various applications by detecting and preventing malicious attacks.

Future Work

The future work drawn from this research is that the proposed ORUSBEM model using BPSO can be utilized as a reliable and effective tool for multi-class intrusion detection in various network environments. This method can be helpful in improving the performance of intrusion detection systems and reduce the number of false positives. Below are the future works of the study:-
(i) While BPSO proved to be effective in selecting the optimal feature subset, other optimization techniques such as Genetic Algorithms (GA), Ant Colony Optimization (ACO), or Differential Evolution (DE) can be used to compare and optimize the results.
(ii) Using the same sampling technique can lead to biased results. Therefore, exploring different sampling techniques such as oversampling, SMOTE or ADASYN could improve the generalization and robustness of the model.
(iii) Hyper-parameters play a crucial role in the model's performance. Therefore, investigating the impact of different hype-rparameters such as the number of estimators in the ensemble, the learning rate or the maximum depth of the tree could improve the model's performance.
(iv) It's essential to validate the performance of intrusion detection methods in real-world network environments, as the dynamics and characteristics of networks can vary widely. Real-world testing can provide insights into the practical applicability of this method.

REFERENCE

Abubakar, M., Surajo, Y. and Tasiu , S. (2025). An Explainable Deep Learning Model for Illegal Dress Code Detection and Classification. *Journal of Basics and Applied Sciences Research*, *3*(1), 1-10. https://dx.doi.org/10.4314/jobasr.v3i1.1

Amru, M., Kannan, R. J., Ganesh, E. N., Muthumarilakshmi, S., Padmanaban, K., Jeyapriya, J., &Murugan, S. (2024). Network intrusion detection system by applying ensemble model for smart home.*International Journal of Electrical and Computer Engineering*, 14(3), 3485–3494. https://doi.org/10.11591/ijece.v14i3, pp3485-3494

Cao, B., Li, C., Song, Y., Qin, Y., Sciences, C. C.-A., & 2022, U. (2022). Network Intrusion Detection Model Based on CNN and GRU. *Mdpi.ComSign In*. https://doi.org/10.3390/app12094184

Imrana S., Obunadike G.N, Abubakar, M. (2025). Machine Learning-Based Framework for Predicting User Satisfaction in E-Learning Systems. *Journal of Basics and Applied Sciences Research*, *3*(2), 78-85. https://dx.doi.org/10.4314/jobasr.v3i2.9

Kumar, V., Das, A. K., & Sinha, D. (2021). UIDS: a unified intrusion detection system for IoT environment. *Evolutionary Intelligence*, *14*(1), 47–59. https://doi.org/10.1007/s12065-019-00291-w

Ieracitano, C., Adeel, A., Gogate, M., Dashtipour, K., Morabito, F. C., Larijani, H., Raza, A., & Hussain, A. (2018). Statistical Analysis Driven Optimized Deep Learning System for Intrusion Detection. *Lecture Notes in* Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10989 LNAI(August), 759–769. https://doi.org/10.1007/978-3-030-00563-4_74

Muhuri, P. S., Chatterjee, P., Yuan, X., Roy, K., & Esterline, A. (2020a). Using a long short-term memory recurrent neural network (LSTM-RNN) to classify network attacks. *Information (Switzerland)*, *11*(5). https://doi.org/10.3390/INFO11050243

Mulyanto, M., Faisal, M., Prakosa, S. W., & Leu, J. S. (2021). Effectiveness of focal loss for minority classification in network intrusion detection systems. *Symmetry*, *13*(1), 1–16. https://doi.org/10.3390/sym13010004

Nagaraja, A., Boregowda, U., Khatatneh, K., Vangipuram, R., Nuvvusetty, R., & Sravan Kiran, V. (2020a). Similarity Based Feature Transformation for Network Anomaly Detection. *IEEE Access*, *8*, 39184–39196. https://doi.org/10.1109/ACCESS.2020.2975716

Qazi, E. U. H., Faheem, M. H., & Zia, T. (2023).HDLNIDS: Hybrid deep-learning-based

network intrusion detection system. *Applied Sciences*, 13(8), 4921.https://doi.org/10.3390/app13084921

Sarumi, O. A., Adetunmbi, A. O., & Adetoye, F. A. (2020a). Discovering computer networks intrusion using data analytics and machine intelligence. *Scientific African*, *9*. https://doi.org/10.1016/j.sciaf.2020.e00500

Sivamohan, S., & Sridhar, S. S. (2023). An optimized model for network intrusion detection systems in industry 4.0 using XAI based Bi-LSTM framework. *Neural Computing and Applications*, 35(15), 11459–11475. https://doi.org/10.1007/s00521-023-08319-0

Tang, C., Luktarhan, N., & Zhao, Y. (2020). An efficient intrusion detection method based on LightGBM and autoencoder. *Symmetry*, *12*(9). https://doi.org/10.3390/sym12091458

Thapa, N., Liu, Z., Kc, D. B., Gokaraju, B., & Roy, K. (2020a). Comparison of machine learning and deep learning models for network intrusion detection systems. *Future Internet*, *12*(10), 1–16. https://doi.org/10.3390/fi12100167

Wu, Z., Wang, J., Hu, L., Zhang, Z., & Wu, H. (2020). A network intrusion detection method based on semantic Re-encoding and deep learning. *Journal of Network and Computer Applications*, *164*. https://doi.org/10.1016/j.jnca.2020.102688

Zhou, Y., Cheng, G., Jiang, S., & Dai, M. (2020a). Building an efficient intrusion detection system based on feature selection and ensemble classifier. *Computer Networks*,*174*. https://doi.org/10.1016/j.comnet.2020.107247