Journal of Computer Science

Research Article

Optimized Feature Reduction Techniques for Enhanced Network Threat Detection in Wireless Sensor Networks

Bikash Kalita and Satyajit Sarmah

Department of Information Technology, Gauhati University, India

Article history
Received: 07-10-2024
Revised: 23-11-2024
Accepted: 13-12-2024

Corresponding Author:
Bikash Kalita
Department of Information Technology, Gauhati University, India
Email: bikax99@gmail.com

Abstract: The security of Wireless Sensor Networks (WSNs) is currently seriously threatened by numerous threats. Consequently, a number of applications are offered to regulate data and information sharing along with the related security features that need to be maintained throughout data transfer. This study suggests an intelligent feature reduction methodology based on machine learning that uses Modified Principal Component Analysis (MPCA) to identify the properties most associated with the attacked classes that are being used. This could help with the machine learning model's complexity. The WSN-DS dataset was used to implement and test the suggested approach. This approach performs very well in intrusion detection for WSNs, attaining great accuracy and dependability. The proposed framework involves three key stages: (1) preprocessing the WSN-DS dataset, (2) applying MPCA to identify and retain the most critical features, and (3) implementing and testing multiple machine learning algorithms including Random forest (RF), Gradient Boosting (GB), Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Neural Networks (NN), on both the original and reduced feature sets. The experiment demonstrates that hybrid feature reduction techniques significantly enhance computational efficiency while maintaining or improving performance, particularly for robust algorithms like RF and GB. RF achieved near-perfect metrics across multiple attack types, with an F-measure of 99.92% for Flooding attacks and an increased recall of 99.70% for Blackhole attacks after reduction. These findings underscore the importance of algorithm selection and feature optimization tailored to specific attack scenarios, establishing hybrid feature reduction as a valuable approach to enhancing threat detection in WSNs.

Keywords: WSN, Network Threats, Feature Reduction, Network Threat Detection

Introduction

Cybercriminals can be prevented from gaining unauthorised access by putting cyber security into place (Behiry & Aly, 2024). The baseline requirements for protecting systems or data from WSN-related threats are confidentiality, integrity, and availability. Keeping up cyber security procedures to protect private data from cybercriminals. Cybersecurity protects cloud services (Zekri *et al.*, 2017), virtual machines, and network topologies. It also assists with forensic investigations and deters cybercrimes. The DNS server needs external protection to prevent hackers from obtaining its data because it is not secure enough.

Machine Learning (ML) is used for the detection and classification of cyberattacks in Wireless Sensor Networks (WSNs) by employing a system that can precisely identify and classify network security threats. Principal Component Analysis can be used to reduce the high-dimensional feature space and improve security breach detection efficiency (More & Mishra, 2020) by extracting and ranking the most pertinent features. The overall goal is to offer efficient features for training and identification systems (Badis *et al.*, 2014a; Nziga & Cannady, 2012) for effective network security threat detection systems.

WSNs have emerged as critical enablers in various domains, including healthcare, environmental monitoring, industrial automation, and military applications. These networks consist of spatially distributed sensors that collect and transmit data to centralized systems. Despite their advantages, WSNs are inherently resource-constrained, characterized by limited computational power, energy resources, and memory.

This makes them highly vulnerable to a range of security threats, including Grayhole, Blackhole, TDMA, and Flooding attacks black hole attack, wormhole attack, sinkhole attack, forwarding attack, (Elsaid & Albatati, 2020)which exploit network vulnerabilities to disrupt communication, compromise data integrity, or exhaust resources (Alsulaiman & Al-Ahmadi, 2021).

Traditional network security threat detection system often relies on high-dimensional datasets, which, while comprehensive, introduce computational overhead and may degrade performance in real-time systems (Harb, 2011). Feature reduction techniques resolve these issues by identifying and retaining the most useful features (Moore *et al.*, 2017), thereby enhancing computational efficiency and potentially improving the performance of machine learning models. This is particularly crucial for WSNs, where processing and energy resources are scarce.

The motivation for selecting hybrid feature reduction techniques stems from their dual advantage: optimizing model performance (Chae *et al.*, 2018) and reducing computational complexity (Amiri *et al.*, 2011). By preserving critical patterns and eliminating redundant features, these techniques enable IDS to focus on the most significant data aspects, ensuring effective detection of anomalous traffic and attacks (Zhao *et al.*, 2020). Moreover, the selection of machine learning algorithms plays a pivotal role, as some models-such as RF and GB - are inherently resilient to feature reduction, making them ideal candidates for resource-constrained environments.

This study evaluates the effectiveness of hybrid feature reduction techniques on the WSN-DS dataset, which is widely used for testing security threat detection systems in WSNs. By comparing multiple machine learning models, including RF, GB, Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and Neural Networks (NN), the research aims to identify the most efficient algorithm-feature combination. The ultimate goal is to establish an approach that ensures robust threat detection while minimizing resource consumption, thereby addressing the practical challenges of securing WSNs

Related Work

Dimensionality Dimensionality reduction and feature selection are crucial steps in enhancing the efficiency of network threat detection systems. Various studies have explored different techniques to achieve this goal. Xiao *et al.* (2007) introduced dynamic tensor analysis (DTA) as a method to handle semi-infinite streams of data, including network traffic data. Tsang *et al.* (2007) proposed “a genetic-fuzzy rule mining approach that also acts as a genetic feature selection wrapper for optimal feature subset search”. Nziga & Cannady (2012) focused on finding the minimal dataset required for successful intrusion detection through dimensionality reduction. Amiri *et al.* (2011) proposed mutual information-based feature selection algorithms for network security threat detection systems, aiming to improve performance compared to existing methods. Moore *et al.* (2017) explored feature extraction and selection using artificial neural networks for Internet threat detection. (Saleh *et al.*, 2019) developed a Hybrid Intrusion Detection System (HIDS) to address the challenge of real-time intrusion detection in the face of large data flows. Prachi *et al.* (2019) evaluated machine learning techniques for network security threats detection and emphasized the importance of feature selection to detect intrusions quickly and accurately. More & Mishra (2020) specifically focused on enhanced-PCA-based feature reduction and selection for real-time network security, highlighting the technique's ability to reduce time complexity and minimize false detection rates. Torabi *et al.* (2021) Conducted a review on feature selection and use of ensemble techniques for anomaly-based network security threats detection system research, emphasizing the effectiveness of these techniques in the training phase and detection process. Finally, Samdekar *et al.* (2021) aimed to enhance the efficiency of intrusion detection in IoT networks through machine learning and bioinspired techniques, focusing on feature selection and dimensionality reduction. Overall, these studies collectively contribute to the development of advanced methods for dimensionality reduction and feature selection in the context of network threat detection.

Several researchers have studied the application of Principal Component Analysis (PCA) for the identification of cyberattacks in computer networks. Wang *et al.* (2004) presented a Principal Component Analysis (PCA) based approach for network-based attack identification. The approach uses the distance between a feature space and its reconstruction onto reduced subspaces to distinguish between various attack types and legitimate activity. Labib & Vemuri (2006) suggested a method for identifying intrusion based on generated statistics and applied PCA to a subset of attack patterns from the Darpa 1998 dataset. A sequence-order-independent technique for network traffic profiling and Application Layer Distributed Danial of Service (DDoS) attack detection was presented by Lee *et al.* (2011). They proposed “a model based on multiple principal component analysis for profiling normal web browsing Behaviours, using reconstruction error as a criterion for detecting DDoS attacks”. Badis *et al.* (2014b) conducted experiments to understand the operational Behaviour of botclouds used for DDoS attacks, highlighting recognizable Behaviour through statistical results based on PCA. Furthermore, they proposed a PCA-based approach for detecting abnormal virtual machine behaviour that could indicate botcloud behaviour supporting DDoS flooding attacks. Fawaz *et al.* (2016) presented an approach for recognising harmful lateral movement with a variety of abnormal host behaviour indicators, such as PCA, and graph-based modelling.

Elnour *et al.* (2020) created an industrial control system security solution using dual isolation forests and PCA pre-processing to identify attacks by distinguishing anomalies. In a recent study by Ullah *et al.* (2021) the authors criticized signature-based Network Intrusion Detection Systems for their inability to identify new attacks and proposed the use of unsupervised learning algorithms based on PCA to detect cybersecurity attacks through the analysis of web browsing activities. The efficacy of PCA is demonstrated by these experiments taken together in detecting various types of cybersecurity attacks in computer networks.

Methodology

This section outlines the proposed methodology for network security threat detection in WSN using ML models. The methodology as shown in Figure 1, involves several stages, each designed to optimize detection accuracy, improve the model's performance, and provide a comprehensive understanding of the security threats in WSNs. The following steps will be followed:

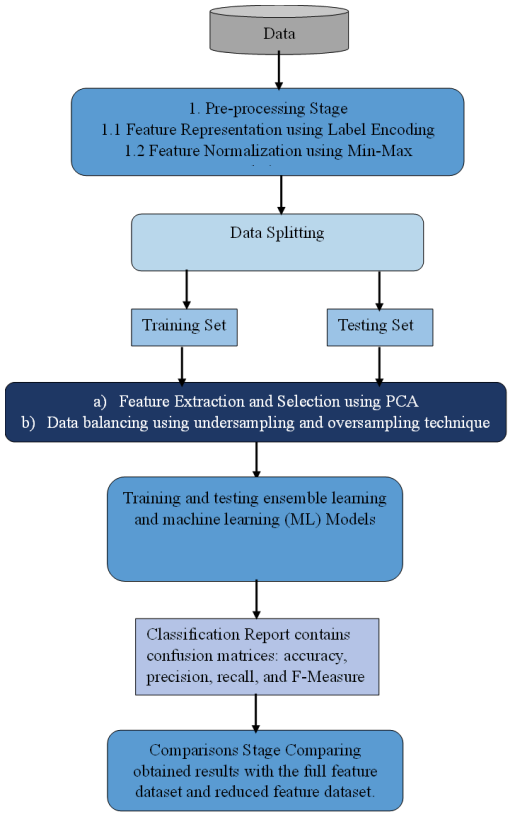

Fig. 1: Proposed architecture workflow for classifying normal and attack traffic in WSN

Dataset Collection and Analysis

The researchers chose the WSN-DS dataset for this study due to its specific focus on Wireless Sensor Networks (WSNs) and its comprehensive representation of various network attack scenarios, making it highly suitable for evaluating network security threat detection techniques in WSNs. As a publicly available dataset, WSN-DS ensures transparency, reproducibility, and comparability with other research in the field. It contains diverse attack types such as Grayhole, Blackhole, TDMA, and Flooding attack, which are critical for testing the effectiveness of threat detection methods. By utilizing this dataset, the authors were able to simulate realistic network conditions and assess their proposed detection mechanisms, while also benefiting from the dataset's established use in the research community for validating results.

The dataset is thoroughly examined to understand the distribution of attack and normal traffic instances. This analysis provides an overview of the data's structure, including the features and their relationships, which helps identify potential challenges like imbalanced classes or missing values. By conducting an initial Exploratory Data Analysis (EDA), the characteristics of the dataset are better understood, ensuring the appropriate application of preprocessing steps in subsequent stages.

Data Preprocessing Stage

The data preprocessing stage plays a pivotal role in enhancing the detection capabilities and performance of the system. Following the proposed methods, this stage encompasses two primary steps: Label Encoding and Feature Normalization.

Label-Based Encoding of Features

Converting non-numeric properties (text or symbols) into numeric values is known as feature encoding. Since intrusion detection datasets frequently include discrete, symbolic, and continuous data, this conversion is crucial. Label encoding and one-hot encoding are the two most used feature encoding methods. Considering the large dimensionality of the dataset, these encoded variables—which are produced for every class—have a substantial impact on how well machine learning approaches work. Thus, label encoding from Scikit-learn is used.

Feature Normalization

Normalization of features is crucial for optimal processing, as it maintains values within the same range. This normalization step ensures uniformity in data processing, facilitating more effective analysis and model training. The Min-Max Scaling method was used to scale the features to a fixed range, This guarantees that every feature makes an equal contribution to the model and aids in enhancing convergence throughout training.

Data Splitting

The dataset is then divided into 70/30 ratios for training and testing sets. By ensuring that the model is tested on unseen data, this phase offers a true assessment of the model's capacity for generalization.

Feature Reduction Strategies

This study employs Principal Component Analysis (PCA) feature reduction methodology to extract and prioritize significant features, thereby addressing dimensionality reduction and enhancing efficiency in cyberattack detection.

Data Balancing Using Undersampling Techniques

To rectify class imbalance, the undersampling technique is used which is a data-balancing strategy that lowers the number of instances in the majority class to equal the size of the minority class. Random undersampling, which chooses a subset of examples at random from the majority class and discards the remainder, is used in this study.

Integration of Machine Learning and Ensemble Learning

The study combines machine learning and ensemble learning to improve detection effectiveness, utilizing each method's strengths to attain higher accuracy. Traditional machine learning algorithms, including DT, NB, KNN, SVM, and NN, are employed alongside ensemble methods like RF and GB. These ensemble methods are particularly effective because they aggregate multiple weak learners to create a more robust and accurate prediction model. The models are tested and compared to evaluate their performance in detecting network security threats, with each algorithm offering distinct advantages in addressing different patterns and complexities within the data.

Comparison: While some researchers may focus exclusively on either ML or DL for intrusion detection, this paper demonstrates the efficacy of blending both methodologies for enhanced outcomes.

Evaluation of Datasets

This study meticulously assesses the proposed methods by subjecting it to the WSN-DS dataset offering a thorough examination of its performance across various scenarios. To evaluate the effect of feature reduction on performance, the evaluation is carried out on both the complete feature dataset and a smaller dataset consisting of 12 selected features. “Is_CH, who CH, ADV_S, ADV_R, JOIN_R, SCH_S, DATA_S, DATA_R, Data_Sent_To_BS, dist_CH_To_BS, send_code, and Expanded Energy” are the specific attributes that are included in the reduced dataset.

Results and Discussion

This section constitutes essential portions in research papers where researchers analyze the study outcomes, offer insights, explanations, and contextualization of their findings. Here are some helpful observations from the experiment, based on the information gained.

Tables 1 to 5, along with their visual representations in Figures 2 to 6 of the experiment reveal that across different attack scenarios and normal traffic, DT, RF, and GB consistently achieve high performance, with only minor differences between the original and reduced feature sets.

Although DT typically retains high accuracy, RF and GB show remarkable stability, especially in recall and F-measure, which makes them reliable for both feature sets. On the other hand, Naive Bayes (NB) exhibits a high degree of sensitivity to feature selection, with notable performance drops when the feature set is reduced.

Table 1: Performance comparison for Normal Traffic detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Algorithms	Original feature set				Reduced to 12 feature set			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure
RF	94.74	94.74	99.45	97.04	94.20	94.20	98.62	96.36
GB	94.33	94.33	98.76	96.5	94.29	94.29	97.93	96.08
DT	94.49	94.49	94.62	94.56	94.33	94.33	94.07	94.20
NB	93.68	93.68	83.86	88.5	93.49	93.49	81.24	86.94
KNN	88.61	88.61	78.34	83.16	88.71	88.71	83.45	86.00
SVM	44.96	44.96	65.24	53.24	60.55	60.55	67.79	50.23
NN	85.19	85.19	59.52	67.12	87.26	87.26	62.10	69.50

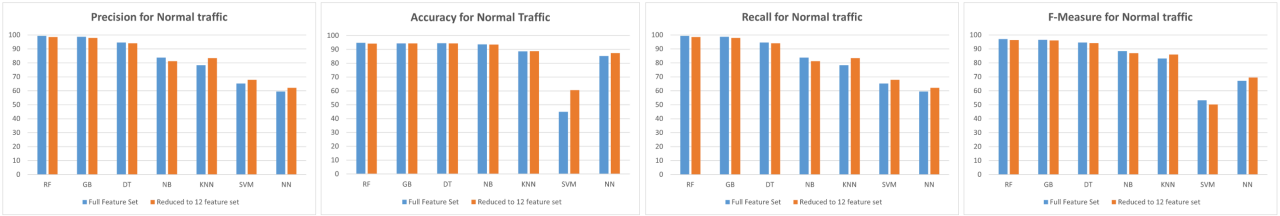

Fig. 2: Performance comparison for Normal Traffic detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

DOI: 10.3844/jcssp.2025.1889.1896

Table 2: Performance comparison for Grayhole Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Algorithms	Original feature set				Reduced to 12 feature set			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure
Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	
RF	98.53	98.53	98.53	98.53	98.51	98.51	97.50	98.00
GB	98.38	98.38	97.94	98.16	97.20	97.20	96.91	97.05
DT	98.81	98.81	98.09	98.45	97.64	97.64	97.21	97.42
NB	72.6	72.6	45.59	56.01	68.14	68.14	86.47	76.22
KNN	65.53	65.53	67.94	66.71	69.14	69.14	67.21	68.16
SVM	48.57	48.57	68.97	49.47	42.91	42.91	60.55	50.23
NN	46.91	46.91	91.62	61.6	51.53	51.53	49.47	48.57

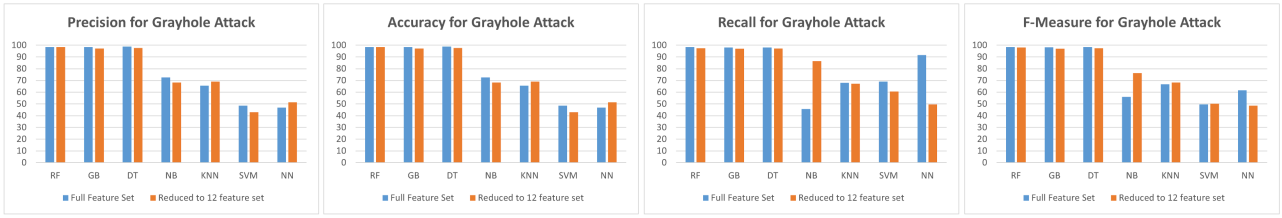

Fig. 3: Performance comparison for Grayhole Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Table 3: Performance comparison for Blackhole Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Algorithms	Original feature set				Reduced to 12 feature set			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure
Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	
RF	98.53	98.53	99.11	98.82	97.54	97.54	99.70	98.61
GB	97.97	97.97	99.85	98.9	96.97	96.97	99.26	98.10
DT	98.82	98.82	98.96	98.89	97.94	97.94	98.67	98.31
NB	49.77	49.77	97.78	65.97	47.53	47.53	98.08	64.03
KNN	70.16	70.16	78.25	73.99	69.22	69.22	72.19	70.67
SVM	58.57	52.57	68.97	59.47	58.42	58.42	67.79	49.04
NN	88.00	88.00	59.62	67.54	68.10	68.10	59.62	67.54

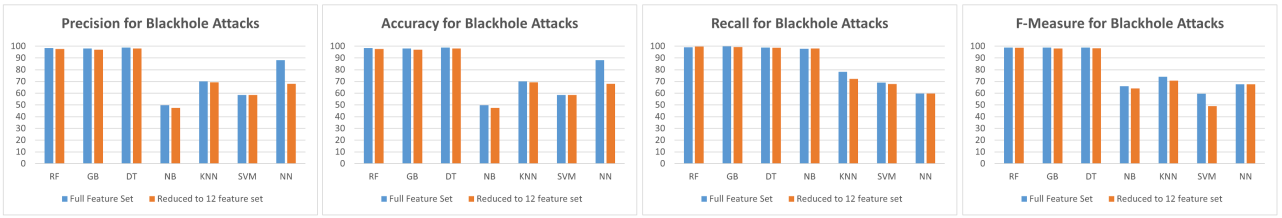

Fig. 4: Performance comparison for Blackhole Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Table 4: Performance comparison for TDMA Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Algorithms	Original feature set				Reduced to 12 feature set			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure
RF	99.54	99.54	94.36	97.1	99.54	99.54	93.64	96.50
GB	99.85	99.85	93.64	96.64	99.08	99.08	93.21	96.05
DT	94.54	94.54	95.09	94.81	94.39	94.39	94.80	94.59
NB	95.83	95.83	33.24	49.36	83.10	83.10	37.79	46.35
KNN	79.88	79.88	74.57	77.13	74.74	74.74	71.82	73.25
SVM	58.57	58.57	68.97	67.47	52.28	52.28	59.27	41.80
NN	88.76	88.76	57.08	63.23	66.78	66.78	73.47	77.90

DOI: 10.3844/jcssp.2025.1889.1896

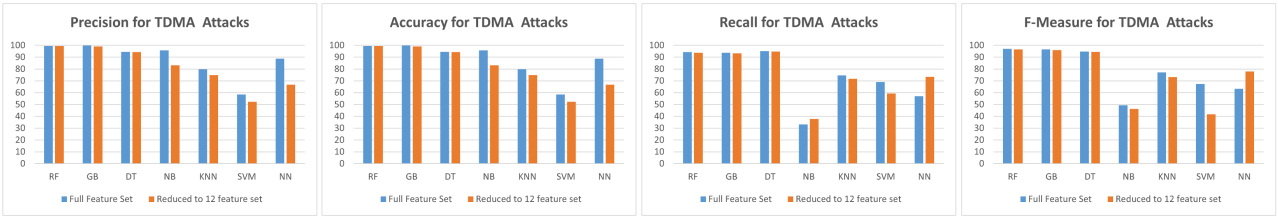

Fig. 5: Performance comparison for TDMA Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Table 5: Performance comparison for Flooding Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

Algorithms	Original feature set				Reduced to 12 feature set			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure
RF	96.64	96.64	99.85	99.92	99.85	99.85	99.85	99.92
GB	95.29	96,92	99.85	99.92	99.91	99.91	99.85	99.92
DT	90.54	90.54	95.85	95.92	99.92	99.92	99.54	99.77
NB	79.67	79.67	95.29	86.78	68.14	68.14	86.47	76.22
KNN	88.76	88.76	92.4	90.54	85.10	85.10	92.86	88.81
SVM	61.90	61.90	56.38	40.75	51.20	51.20	56.38	40.75
NN	66.78	66.78	93.47	77.9	88.42	88.42	66.03	67.85

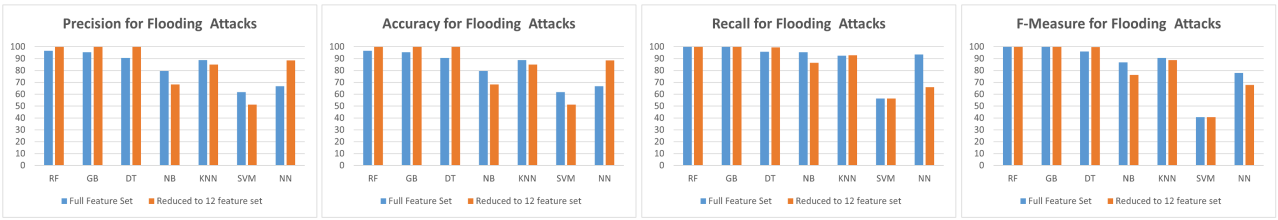

Fig. 6: Performance comparison for Flooding Attack detection using the full feature set versus a reduced 12-feature set from the WSN-DS dataset, showing accuracy, precision, recall, and F-measure

SVM and NN exhibit significant variability in performance across different datasets and feature sets. SVM shows considerable improvement with feature reduction in some cases but remains inconsistent overall. Neural Networks, although showing some improvement with feature reduction, generally perform poorly compared to other algorithms. K-Nearest Neighbors (KNN) shows a slight preference for the original feature set but with marginal differences, indicating moderate robustness.

In conclusion, feature reduction has a mixed impact, enhancing some algorithms while slightly degrading others. RF and GB stand out for their robustness across different scenarios, maintaining high performance regardless of the feature set size. The performance variability of NB, SVM, and NN suggests that these algorithms may require further tuning or more complex architectures for consistent results. Overall, the choice of algorithm and the decision to reduce features should be tailored to the specific attack type and dataset to achieve optimal performance.

Conclusion

This study demonstrates the significant advantages of using hybrid feature reduction techniques for network threat detection in wireless sensor networks (WSN). This study comprehensively evaluated the performance of various machine learning algorithms on the WSN-DS dataset for detecting normal traffic and specific attack types (Grayhole, Blackhole, TDMA, and Flooding attacks). Key findings reveal that the RF algorithm consistently demonstrated superior performance across most scenarios, achieving high accuracy, precision, recall, and F-measure. GB followed closely, showing comparable results, particularly in scenarios involving Flooding and TDMA attacks.

Future studies will be focused on the following strategies:

Hybrid Approaches: Develop ensemble models that combine the strengths of algorithms like RF and GB with complementary methods to enhance detection robustness and address weaknesses in recall and precision.

Advanced Feature Engineering: Explore automated feature selection techniques such as genetic algorithms or deep learning-based embeddings to improve model performance with reduced feature sets.

Attack-Specific Optimizations: Tailor machine learning models to individual attack profiles, leveraging

domain knowledge to optimize detection capabilities for complex or less frequent attacks.

Scalability and Real-Time Implementation: Evaluate the algorithms in real-world deployments to test scalability and responsiveness in dynamic WSN environments.

Acknowledgment

The authors of this manuscript would like to express their appreciation and gratitude to their universities for supporting this research.

The authors would like to thank the editors for their efforts in handling the manuscript and all reviewers for the constructive comments which improved the original submission.

Author's Contributions

Bikash Kalita: Acquisition of data, investigation, software, original draft preparation, approved the version to be submitted and any revised version.

Satyajit Sarmah: Analysis and interpretation of data, review & editing, approved the version to be submitted and any revised version.

Ethics

This study is original and innovative and contains unpublished material. The corresponding author confirms that all the other authors have read and approved the manuscript and no ethical issues involved or conflicts of interest to release.

References

Alsulaiman, L., & Al-Ahmadi, S. (2021). Performance Evaluation of Machine Learning Techniques for DOS Detection in Wireless Sensor Network. *International Journal of Network Security & Its Applications*, *13*(2), 21–29. https://doi.org/10.5121/ijnsa.2021.13202

Amiri, F., Rezaei Yousefi, M., Lucas, C., Shakery, A., & Yazdani, N. (2011). Mutual Information-Based Feature Selection for Intrusion Detection Systems. *Journal of Network and Computer Applications*, *34*(4), 1184–1199. https://doi.org/10.1016/j.jnca.2011.01.002

Badis, H., Doyen, G., & Khatoun, R. (2014a). Toward a Source Detection of Botclouds: A PCA-Based Approach. *Monitoring and Securing Virtualized Networks and Services*, *8508*, 105–117. https://doi.org/10.1007/978-3-662-43862-6_13

Badis, H., Doyen, G., & Khatoun, R. (2014b). Understanding Botclouds From a System Perspective: A Principal Component Analysis. *2014 IEEE Network Operations and Management Symposium (NOMS)*, 1–9. https://doi.org/10.1109/noms.2014.6838310

Behiry, M. H., & Aly, M. (2024). Cyberattack detection in wireless sensor networks using a hybrid feature reduction technique with AI and machine learning methods. *Journal of Big Data*, *11*(1), 1–23. https://doi.org/10.1186/s40537-023-00870-w

Chae, H., Jo, B., Choi, S.-H., & Park, T. (2018). Feature Selection for Intrusion Detection using NSL-KDD. *Recent Advances in Computer Science*, 184–187.

Elnour, M., Meskin, N., Khan, K., & Jain, R. (2020). A Dual-Isolation-Forests-Based Attack Detection Framework for Industrial Control Systems. *IEEE Access*, *8*, 36639–36651. https://doi.org/10.1109/access.2020.2975066

Elsaid, S. A., & Albatati, N. S. (2020). An optimized collaborative intrusion detection system for wireless sensor networks. *Soft Computing*, *24*(16), 12553–12567. https://doi.org/10.1007/s00500-020-04695-0

Fawaz, A., Bohara, A., Cheh, C., & Sanders, W. H. (2016). Lateral Movement Detection Using Distributed Data Fusion. *2016 IEEE 35th Symposium on Reliable Distributed Systems (SRDS)*, 21–30. https://doi.org/10.1109/srds.2016.014

Harb, H. M. (2011). Selecting Optimal Subset of Features for Intrusion Detection Systems. *Advances in Computational Sciences and Technology*, *4*(2), 179–192.

Moore, K. L., Bihl, T. J., Bauer, K. W., & Dube, T. E. (2017). Feature Extraction and Feature Selection for Classifying Cyber Traffic Threats. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, *14*(3), 217–231. https://doi.org/10.1177/1548512916664032

More, P., & Mishra, P. (2020). Enhanced-PCA Based Dimensionality Reduction and Feature Selection for Real-Time Network Threat Detection. *Engineering, Technology & Applied Science Research*, *10*(5), 6270–6275. https://doi.org/10.48084/etasr.3801

Nziga, J.-P., & Cannady, J. (2012). Minimal Dataset for Network Intrusion Detection Systems via MID-PCA: A Hybrid Approach. *2012 6th IEEE International Conference Intelligent Systems*, 453–460. https://doi.org/10.1109/is.2012.6335176

Prachi, Malhotra, H., & Sharma, P. (2019). Intrusion Detection using Machine Learning and Feature Selection. *International Journal of Computer Network and Information Security*, *11*(4), 43–52. https://doi.org/10.5815/ijcnis.2019.04.06

Saleh, A. I., Talaat, F. M., & Labib, L. M. (2019). A Hybrid Intrusion Detection System (HIDS) Based on Prioritized k-Nearest Neighbors and Optimized SVM Classifiers. *Artificial Intelligence Review*, *51*(3), 403–443. https://doi.org/10.1007/s10462-017-9567-1

Samdekar, R., Ghosh, S. M., & Srinivas, K. (2021). Efficiency Enhancement of Intrusion Detection in Iot Based on Machine Learning Through Bioinspire. *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 383–387. https://doi.org/10.1109/icicv50876.2021.9388392

Torabi, M., Udzir, N. I., Abdullah, M. T., & Yaakob, R. (2021). A Review on Feature Selection and Ensemble Techniques for Intrusion Detection System. *International Journal of Advanced Computer Science and Applications*, *12*(5), 470–477. https://doi.org/10.14569/ijacsa.2021.0120566

Tsang, C.-H., Kwong, S., & Wang, H. (2007). Genetic-Fuzzy Rule Mining Approach and Evaluation of Feature Selection Techniques for Anomaly Intrusion Detection. *Pattern Recognition*, *40*(9), 2373–2391. https://doi.org/10.1016/j.patcog.2006.12.009

Ullah, I., Mengersen, K., Hyndman, R. J., & McGree, J. (2021). Detection of cybersecurity attacks through analysis of web browsing activities using principal component analysis. *Cryptography and Security (Cs.CR); Methodology (Stat.ME)*. http://arxiv.org/abs/2107.12592

Xiao, Y., Rayi, V. K., Sun, B., Du, X., Hu, F., & Galloway, M. (2007). A Survey of key Management Schemes in Wireless Sensor Networks. *Computer Communications*, *30*(11–12), 2314–2341. https://doi.org/10.1016/j.comcom.2007.04.009

Zekri, M., Kafhali, S. E., Aboutabit, N., & Saadi, Y. (2017). DDoS attack detection using machine learning techniques in cloud computing environments. *2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech)*, 1–7. https://doi.org/10.1109/cloudtech.2017.8284731