A lightweight machine learning approach for denial-of-service attacks detection in wireless sensor networks

Mohamed Loughmari, Anass El Affar
Engineering Sciences Laboratory (LSI), Polydisciplinary Faculty of Taza, Sidi Mohamed Ben Abdellah University, Taza, Morocco

Article Info

Article history:

Received Jul 17, 2024
Revised Sep 3, 2024
Accepted Oct 1, 2024

Keywords:

Cybersecurity
Denial of service
Intrusion detection
Machine learning
Wireless sensor network
Wireless sensor network security

ABSTRACT

Wireless sensor networks (WSNs) are increasingly prevalent in the Internet of Things ecosystem and have been used in several fields such as environmental monitoring, military, and healthcare. However, their limited resources and distributed architecture remain two main challenges: energy and security. Furthermore, denial of service (DoS) attacks are one of the principal cyber threats to WSNs. This research proposes a lightweight machine learning (ML) approach based on the extreme gradient boosting (XGBoost) model to detect these attacks in WSNs. Through an extensive investigation, we evaluate four prominent ML algorithms: random forest (RF), k-nearest neighbor (KNN), stochastic gradient descent (SGD), and XGBoost, using the WSN-DS dataset. In addition, we implement and investigate several feature selection techniques in order to have an improved version of the original dataset. Moreover, we evaluate the performance using various performance metrics, which include accuracy, precision, recall, F1-score, and processing time. The latter is a crucial consideration in WSN environments. For validation, we have employed 5-fold cross-validation to ensure robust and reliable results. The proposed model has achieved good performance in all metrics, with a maximum accuracy of up to 99.73%, and a 68% lower processing time compared to the other investigated classifiers.

Corresponding Author:

Mohamed Loughmari
Engineering Sciences Laboratory (LSI), Polydisciplinary Faculty of Taza, Sidi Mohamed Ben Abdellah University, B.P. 1223, Taza, Morocco
Email: mohamed.loughmari@usmba.ac.ma

1. INTRODUCTION

Wireless sensor networks (WSNs) are becoming an essential piece of the internet of things (IoT) ecosystem, an integral part of modern technology, and have been adopted for various application sectors such as environmental monitoring, military, healthcare, agriculture [1]. These wireless networks consist of many autonomous cheap and low-power sensor nodes, which are deployed across various areas to collect important data and then transmitted in a cooperative manner via wireless means to a higher performance node commonly referred to as the sink node or base station [2]. However, due to their wireless communication nature, limited computational resources, and energy constraints, WSNs are exposed to various cyber-attacks, especially denial of service (DoS) attacks [3]. DoS attack aims to compromise the availability and integrity of the WSN systems, by overwhelming the sensor nodes, causing service disruptions and data loss. protecting against DoS attacks is very challenging because they can be generated from anywhere, particularly in the context of wireless sensors deployed on a massive scale. Current network security solutions, such as traditional intrusion detection systems (IDS) may not be sufficient for WSNs as they differ from usual computer networks, the prevention protection technology designed for WSN infrastructure must be highly accurate, fast, and suitable to ensure minimal overhead [4], [5]. Therefore, there is a pressing need for an efficient lightweight intrusion detection technique.

WSNs are, like any network type, liable to various security threats and vulnerabilities. However, their unique characteristics, like resource constraints and deployment environments, require specialized defense schemes. Numerous existing studies addressed these challenges by investigating the potential of machine learning (ML) models for cyber threat detection while also highlighting the difficulties due to resource constraints [6]. Ravi *et al.* [7] presented an end-to-end deep learning (DL) approach using recurrent neural networks (RNNs) for intelligent network intrusion detection and classification. They evaluated their work mainly on the software-defined networking-IoT (SDN-IoT) [8] and various benchmark datasets, including KDD-Cup-1999 [9], UNSW-NB15 [10], WSN-DS [2], and CICIDS-2017 [11], achieving a maximum accuracy of 99% for intrusion detection and 97% for intrusion classification. A sophisticated intrusion detection method (SLGBM) specifically designed for WSNs was proposed by Jiang *et al.* [12], which combines the sequential backward selection (SBS) feature selection technique and the LightGBM algorithm, their experimental results detection rates exceeded 96% for each of the five categories within the WSN-DS dataset. In their work, Wazirali and Ahmad [13] investigated a variety of ML algorithms to analyze WSN traffic data and identify DoS/DDoS attacks, and among the tested algorithms, GBoost achieved the highest average performance, with 99.6% accuracy and a moderate training time of 1.41 seconds. Ramana *et al.* [14] developed a novel intrusion detection system (WOGRU-IDS) that employed whale optimization algorithm (WOA) to optimize the hyperparameters of a long short-term memory (LSTM) network. Their results demonstrated an average accuracy of 99.85% for all attack types. Meenakshi and Karunkuzhali [15] proposed a novel intrusion detection method employing color wiener filtering (CWF) in the preprocessing phase and Tasmanian devil optimization (TDO) in the feature selection phase while utilizing optimized self-attention-based provisional variational auto-encoder generative adversarial network (SAPVAGAN) for binary classification of WSN-DS data, their novel method provides higher accuracy and lower computation time when compared solely to SLGBM [12], RNN [7], and WOGRU [14] methods. Manjula and Priya [16] introduced a VGG-19-based CNN-LSTM framework to achieve a high classification accuracy of 98.86%. Liu *et al.* [17] suggested an edge-based intrusion detection model (KNN-PL-AOA) tailored for WSNs, their approach employed an improved k-nearest neighbor (KNN) classifier integrated with a novel optimization algorithm (PL-AOA), the model demonstrated promising performance with an accuracy rate of 99%. Sivagaminathan *et al.* [18] examined the application of particle swarm optimization (PSO) as a feature selection technique in conjunctions with various ML algorithms, the results suggested that this combination achieved superior classification accuracy compared to other methods in their case study. Rameshkumar *et al.* [19] assayed reinforcement learning (RL) by combining a deep Q network with transfer learning (TLDQN) approach for DDoS nature attacks in wireless multimedia sensor networks (WMSNs) for smart agriculture applications. The authors claimed that their technique achieves 90% throughput, 95% packet delivery ratio, 89% DDoS detection accuracy, and 95% multipath analysis accuracy. In 2024, Shakya *et al.* [20] highlighted the potential of combining DL and RL for intrusion detection in WSNs by proposing an innovative approach: IRADA integrated RL-based advanced DL. The evaluation showed impressive results with high accuracy (99.50%), specificity (99.94%), sensitivity (99.48%), F1-score (98.26%), Kappa statistics (99.42%), and area under the curve (AUC) (99.38%).

Despite the extensive research efforts conducted in the field of intrusion detection, developing an applicable solution with the ability to effectively detect DoS threats in WSN environments remains a significant challenge. This work presents a novel approach for detecting DoS attacks in WSNs. The main contributions of this paper are: A lightweight approach leveraging ensemble boosting learning techniques: that aims to provide an efficacious and efficient solution for securing WSNs against DoS attacks and potential intrusions by offering high accuracy, robust feature selection, comprehensive evaluation metrics, and reduced processing time making it a suitable solution for real-time detection in such environments; A novel IDS model based on extreme gradient boosting (XGBoost), and the utilization of the recent, comprehensive WSN-DS dataset, which is designed specifically for evaluating DoS attacks in WSNs, moreover, The work is rigorously evaluated through cross-validation to ensure that our findings are both relevant and applicable to real-world scenarios in WSN environments. The rest of the paper is structured as follows: section 2 describes the proposed approach in detail, section 3 illustrates the results and discussion, and section 4 concludes this manuscript.

2. RESEARCH METHOD

This section introduces the proposed methodology; in the first stage we begin with the data collection of WSN-DS, a wireless sensor network specialized dataset, developed to help for better detection and classification of DoS attacks, it is the dataset on which we train and test the performance of ML models. In the second stage, we preprocess the collected dataset by handling missing and null values, besides performing data encoding. In the third stage, we examine feature selection techniques where we aim to only keep significant features, thereby enhancing performance and minimizing computational overhead, In the fourth stage, we implement four variant machine learning algorithms: extreme gradient boosting (XGBoost), random forest (RF), k-nearest neighbor (KNN), and stochastic gradient descent (SGD); That are selected for their significant impact and contributions to defending against security attacks [21], specifically their ability in intrusion detection. Finally, we evaluate our proposed methodology by performing cross-validation after adjusting the evaluation metrics of the investigated ML models, and proposing the best-performing model for efficient DoS detection in WSN environments. The schematic diagram in Figure 1 illustrates the proposed system.

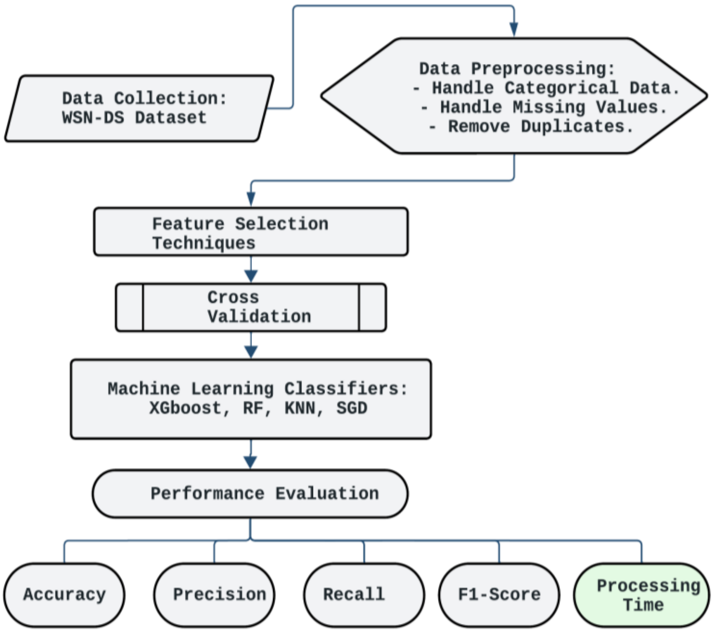

Figure 1. Schematic diagram of the proposed system

2.1. Dataset overview

This paper utilizes a synthetic dataset constructed especially for denial-of-service detection in WSNs by Almomani *et al.* [2], named WSN-DS. The data were collected using the LEACH protocol [22] and network simulator NS-2; 23 features were extracted, but only 19 were presented in the dataset, including the target. Table 1 summarizes these features and corresponding descriptions.

Table 1. Features of the WSN-DS dataset

Id	Feature name	Description
1	id	Unique identifier for a specific sensor node
2	Time	The runtime at which the sensor node data was captured
3	Is_CH	A flag indicating the node's role (Cluster head or regular sensor node)
4	who CH	Unique identifier of the cluster head (CH)
5	Dist_To_CH	physical distance between the sensor node and the assigned CH
6	ADV_S	The number of advertisement CH messages broadcasted to nodes
7	ADV_R	The number of advertisement messages received from CH
8	JOIN_S	The number of join request messages sent by the node to a chosen CH
9	JOIN_R	The number of join request messages received by the CH from sensor nodes
10	SCH_S	The number of messages broadcasted by the CH containing the TDMA schedule
11	SCH_R	The number of messages received by the node containing the TDMA schedule
12	Rank	The assigned order in the node TDMA scheduling
13	DATA_S	The number of data packets successfully sent from the node to its CH
14	DATA_R	The number of data packets received by the node from its assigned Cluster CH
15	Data_Sent_To_BS	The number of data packets successfully transmitted by the node to BS
16	dist_CH_To_BS	The distance between CH and BS
17	send_code	A code identifying the cluster to which the node belongs
18	Expaned energy	amount of energy consumed by the sensor node
19	Attack type	type of node behavior observed (Normal or malicious behavior)

WSN-DS has 374,661 records representing normal and anomalous traffic in four DoS attack types: blackhole, grayhole, TDMA, and flooding. Figure 2 shows a graphical representation of the dataset, displaying the distribution of all five types of attacks.

a. Blackhole attack: this is when a compromised node behaves as the best route for all incoming traffic, then discards and drops all data packets from reaching their intended destinations.
b. Grayhole attack: unlike the first one, gray hole did not perform a complete blockage but randomly or selectively dropped some data packets.
c. Time division multiple access (TDMA) or scheduling: this attack exploits the scheduled access of TDMA by manipulating this protocol to disrupt the scheduling and coordination of transmissions between nodes, causing collisions, delays, or interference.
d. Flooding: this attack aims to overwhelm the network with a large volume of traffic, which can lead to denial of service for legitimate users or nodes by consuming infinite energy, network traffic, and memory.

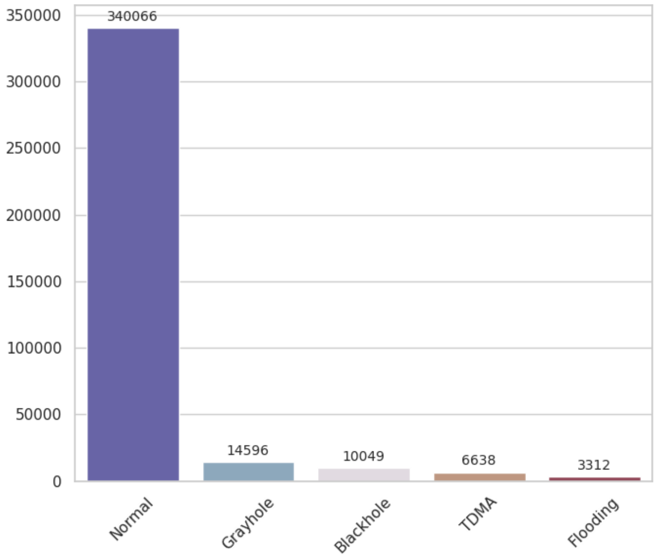

Figure 2. Distribution of WSN-DS dataset

2.2. Data preprocessing

The WSN-DS target feature consists of strings that represent different attack types. These textual labels are translated into numerical representations to prepare and facilitate uniform data handling for algorithmic processing. Normal traffic is encoded as 0, while the four attack types: blackhole, grayhole, flooding, and TDMA, are encoded as 1, 2, 3, and 4, respectively. The other data preprocessing operations include identifying and handling missing values, null values, infinity values, and duplicate removal in order to preserve data fidelity and mitigate potential sources of bias or noise.

2.3. Feature selection

The feature selection technique is intended to select the optimal features in the dataset. It is a powerful phase of any machine learning process, and it typically impacts the classifier's performance, data understanding, resources, and time consumption [23]. This is particularly important in the context of the research, where we are dealing with a WSN environment that often has limited computing resources.

Therefore, this research explored various feature selection techniques, such as recursive feature elimination (RFE), sequential backward selection and forward (SBSF), and feature importance-based methods. However, these approaches did not yield the desired improvements in the results. The main progress was achieved through a straightforward yet efficient action - the exclusion of the index column, which was determined to be insignificant for the detection process. Furthermore, from the generated SNS heatmap of feature correlation in Figure 3, we dropped the ‘SCH_R’ feature based on the analysis to identify and remove the highly correlated features, this was done in preference to deleting the ‘JOIN_S’ feature, as it was deemed to have greater feature importance. In addition, the absence of any columns with a single unique value was checked and confirmed. After this iterative feature selection process, we were left with 16 features from the original dataset, which we then used to train and evaluate the examined models.

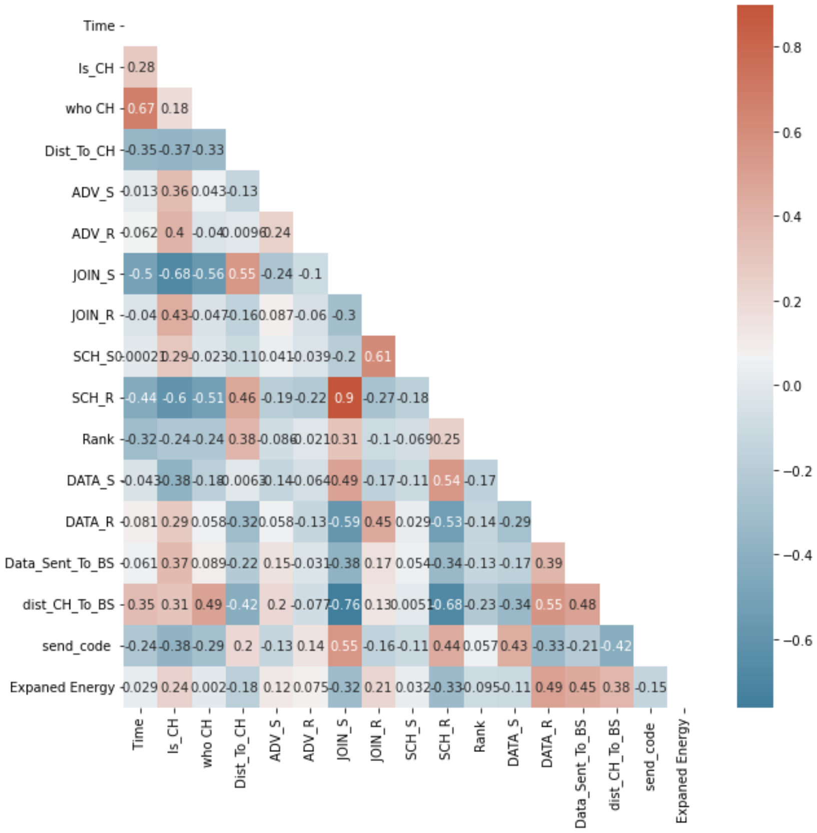

Figure 3. Heatmap of feature correlation

2.4. Validation

A cross-validation was performed to validate our proposed model reasonably and objectively. We employed 5 folds to evaluate its performance on unseen data. Furthermore, the suggested model was evaluated using several performance metrics, including accuracy, precision, recall, F1-score, and processing time. The calculation of these measures relies on the values from the confusion matrix, including true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

a. Accuracy: consider both instances of normal and abnormal data with other instances.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{1}$$

b. Precision: percentage of correctly classified abnormal (attack) instances out of all instances classified as abnormal.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

c. Recall: proportion of correctly classified abnormal samples relative to the total number of actual abnormal samples.

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

d. F1-score: Combination of precision and recall into a single score value by taking their harmonic mean.

$$F1 - score = \frac{2\ x\ Precision\ x\ Recall}{Precision+Recall} \tag{4}$$

2.5. Experimental environment

The selection of the hyperparameter combination for the investigated ML models was optimized using the randomized search technique [24] from the RandomizedSearchCV scikit-learn library, which efficiently explored a wide range of hyperparameter values, to ensure that the models were fine-tuned for the best performance. The experiments presented in this paper were conducted on the Google Colab platform, a cloud-based Jupyter Notebook environment. The runtime was specifically as follows:

- Runtime Type: Python 3
- Hardware Accelerator: CPU
- RAM: 12.7 GB

3. RESULTS AND DISCUSSION

Our research methodology evaluates our proposed approach in both original and improved datasets after the feature selection phase; All models are trained and tested in this phase, making a clear classification report. As mentioned before in the improved version of the dataset we end up with 16 features. The feature selection techniques applied in this study did not lead to a significant improvement and often led to decreased performance, which indicates that 18 features present in the original publicly available dataset were carefully selected from the originally reported 23 features and are crucial for WSN security purposes [2].

Figures 4 and 5 expose the Accuracy performance for the original and the improved versions of the dataset, respectively, as we can observe that XGBoost and RF classification models achieved and maintained high accuracy, exceeding 99.7% for both datasets with no benefit from the dataset improvements. KNN shows a slight improvement in the other dataset version. The most significant change was observed in SGD, where accuracy decreased by nearly 23 percentage points. Overall, the results highlight the robustness of ensemble methods like XGBoost and RF.

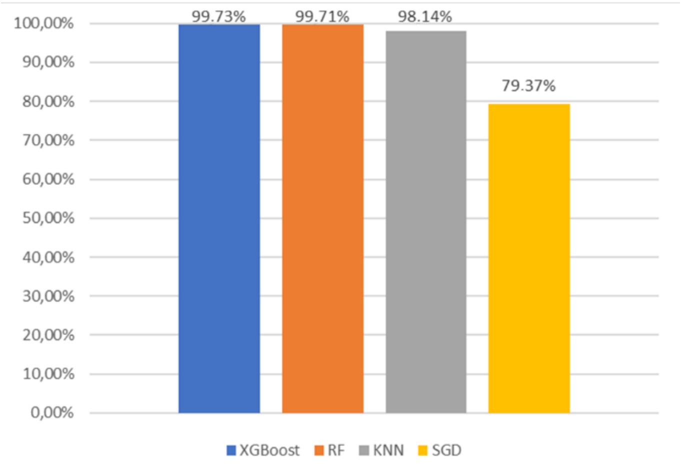

Figure 4. Accuracy exhibited by all classifiers on the original dataset

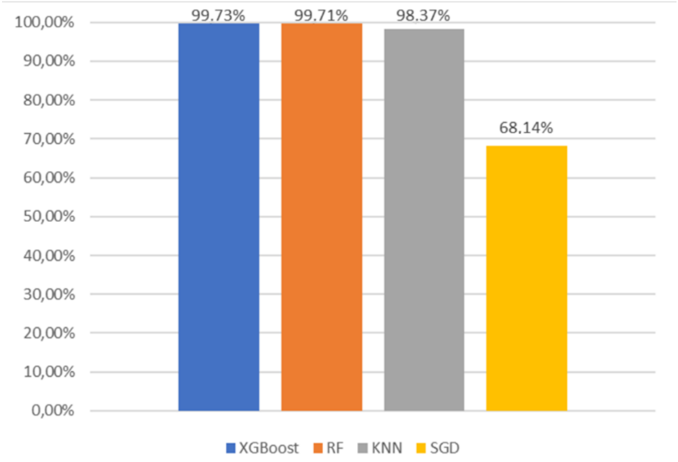

Figure 5. Accuracy exhibited by all classifiers on the improved dataset

Tables 2 and 3 report detailed performance of all models using 5-fold cross-validation on the original and improved datasets, respectively. The metrics reported include accuracy, precision, recall, F1-score, and time consumption. Results demonstrate that XGBoost and RF perform consistently well in both dataset versions by maintaining high precision, recall, and F1-score values. In terms of time processing XGBoost is the fastest model and is more rapid in the enhanced version of the dataset. KNN is the highest, and RF/SGD are also relatively high, with slight time enhancement in the improved dataset version.

Table 2. Performance metrics of the original dataset

Classifiers	Performance metrics				Time consumption (s)
	Accuracy	Precision	Recall	F1-score	
XGBoost	99.7328%	99.7334%	99.7328%	99.7306%	85.46113570699998
RF	99.7074%	99.7098%	99.7074%	99.7058%	200.41408549500005
KNN	98.1433%	98.0852%	98.1433%	98.0903%	572.8316935019999
SGD	79.3685%	84.0467%	79.3685%	78.1847%	273.686294588

Table 3. Performance metrics of the improved dataset

Classifiers	Performance metrics				Time consumption (s)
	Accuracy	Precision	Recall	F1-score	
XGBoost	99.7320%	99.7326%	99.7320%	99.7298%	83.923988580000
RF	99.7109%	99.7129%	99.7109%	99.7093%	193.997570229000
KNN	98.3665%	98.3208%	98.3665%	98.3267%	555.245838744000
SGD	68.1439%	86.0112%	68.1439%	67.6183%	263.278639166999

Comparing XGBoost to RF indicates that both perform well and consistently in both versions of the dataset confirming that ensemble learning techniques are effective on large-scale data [25]. XGBoost offers slightly better performance, which is notable in the context of cybersecurity, where even small differences in rates can have significant implications in real-world scenarios. However, RF has a considerable processing time which raises concerns about computational consumption and may render it unsuitable for a limited resource environment. KNN also gives respectable results performance but lags behind XGBoost and RF in all performance metrics. In addition, it has the highest time consumption. SGD is not a reliable classifier for the intrusion detection task since it performs the worst among all the classifiers in terms of all metrics.

The key findings indicate that our proposed XGBoost model demonstrates the highest performance in terms of all metrics. Moreover, it has significantly lower processing time compared to RF, KNN, and SGD. More precisely, for the original dataset version, XGBoost's processing time is approximately 57.36% lower than RF, 85.06% lower than KNN, and 68.78% lower than SGD, for the second dataset version, the decreases are approximately 56.73%, 84.88%, and 68.12%, respectively. These results indicate that the lightweight developed model is applicable for limited resource networks like the WSN, where energy efficiency is critical, and any developed IDS must consume minimal energy. Limitations include real-time implementation, scalability concerns, and dataset constraints which may not fully represent the wide and recent variety of DoS attacks. In order to demonstrate the effectiveness and efficiency of the suggested model, we conducted a comparison with other relevant papers, as shown in Figure 6.

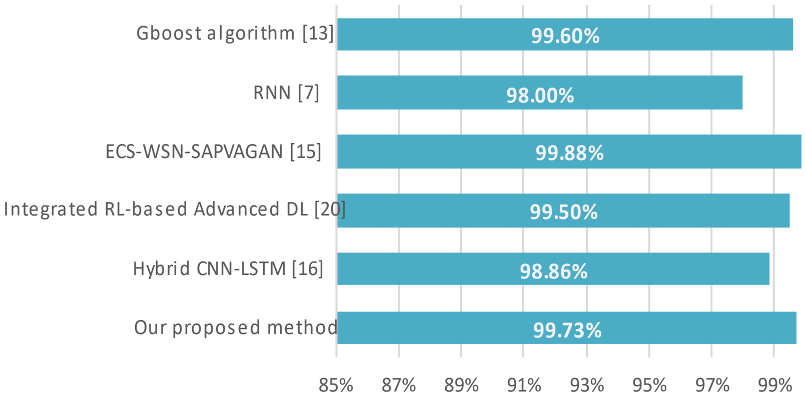

Figure 6. Comparison of our proposed method and other methods on WSN-DS

4. CONCLUSION

Detecting denial of service in WSNs based on intrusion detection with high accuracy and low computing consumption remains a significant challenge task; In this study, we have demonstrated a lightweight and effective approach tailored for this purpose. Through extensive experimentation and comparison conducted, we found that XGBoost provides high accuracy up to 99.73%, and reduces processing time by 68% when detecting malicious activities, compared to random forest (RF), k-nearest neighbor (KNN), and stochastic gradient descent (SGD) models; and we conclude that XGBoost is the most suitable machine learning classifier for WSN constraints since it achieves the highest accuracy and the lowest processing time which are a highly important factors in addition to the other evaluation metrics. In future work, we plan to consider the integration of the proposed model in real-time detection, to examine its effectiveness on real data generated from IoT devices in operational settings.

REFERENCES

[1] M. A. Jamshed, K. Ali, Q. H. Abbasi, M. A. Imran, and M. Ur-Rehman, “Challenges, applications, and future of wireless sensors in internet of things: a review,” *IEEE Sensors Journal*, vol. 22, no. 6, pp. 5482–5494, Mar. 2022, doi: 10.1109/JSEN.2022.3148128.

[2] I. Almomani, B. Al-Kasasbeh, and M. Al-Akhras, “WSN-DS: a dataset for intrusion detection systems in wireless sensor networks,” *Journal of Sensors*, vol. 2016, pp. 1–16, 2016, doi: 10.1155/2016/4731953.

[3] B. Bhushan and G. Sahoo, “Recent advances in attacks, technical challenges, vulnerabilities and their countermeasures in wireless sensor networks,” *Wireless Personal Communications*, vol. 98, no. 2, pp. 2037–2077, Sep. 2018, doi: 10.1007/s11277-017-4962-0.

[4] A. Mitrokotsa and A. Karygiannis, “Intrusion detection techniques in sensor networks,” *Network Security*, pp. 251–272, 2008.

[5] I. Batra, S. Verma, Kavita, and M. Alazab, “A lightweight IoT-based security framework for inventory automation using wireless sensor network,” *International Journal of Communication Systems*, vol. 33, no. 4, Oct. 2020, doi: 10.1002/dac.4228.

[6] K. Shaukat *et al.*, “Performance comparison and current challenges of using machine learning techniques in cybersecurity,” *Energies*, vol. 13, no. 10, May 2020, doi: 10.3390/en13102509.

[7] V. Ravi, R. Chaganti, and M. Alazab, “Recurrent deep learning-based feature fusion ensemble meta-classifier approach for intelligent network intrusion detection system,” *Computers and Electrical Engineering*, vol. 102, Sep. 2022, doi: 10.1016/j.compeleceng.2022.108156.

[8] A. Kaan Sarica and P. Angin, “A novel SDN dataset for intrusion detection in IoT networks,” in *16th International Conference on Network and Service Management, CNSM 2020, 2nd International Workshop on Analytics for Service and Application Management, AnServApp 2020 and 1st International Workshop on the Future Evolution of Internet Protocols, IPFutu*, Nov. 2020, pp. 1–5, doi: 10.23919/CNSM50824.2020.9269042.

[9] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “A detailed analysis of the KDD CUP 99 data set,” in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, Jul. 2009, pp. 1–6, doi: 10.1109/CISDA.2009.5356528.

[10] N. Moustafa and J. Slay, “The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set,” *Information Security Journal*, vol. 25, no. 1–3, pp. 18–31, Jan. 2016, doi: 10.1080/19393555.2015.1125974.

[11] I. Sharafaldin, A. Habibi Lashkari, and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” in *Proceedings of the 4th International Conference on Information Systems Security and Privacy*, 2018, pp. 108–116, doi: 10.5220/0006639801080116.

[12] S. Jiang, J. Zhao, and X. Xu, “SLGBM: an intrusion detection mechanism for wireless sensor networks in smart environments,” *IEEE Access*, vol. 8, pp. 169548–169558, 2020, doi: 10.1109/ACCESS.2020.3024219.

[13] R. Wazirali and R. Ahmad, “Machine learning approaches to detect DoS and their effect on WSNs lifetime,” *Computers, Materials and Continua*, vol. 70, no. 3, pp. 4921–4946, 2022, doi: 10.32604/cmc.2022.020044.

[14] K. Ramana, A. Revathi, A. Gayathri, R. H. Jhaveri, C. V. L. Narayana, and B. N. Kumar, “WOGRU-IDS — An intelligent intrusion detection system for IoT assisted wireless sensor networks,” *Computer Communications*, vol. 196, pp. 195–206, Dec. 2022, doi: 10.1016/j.comcom.2022.10.001.

[15] B. Meenakshi and D. Karunkuzhali, “Enhancing cyber security in WSN using optimized self-attention-based provisional variational auto-encoder generative adversarial network,” *Computer Standards and Interfaces*, vol. 88, Mar. 2024, doi: 10.1016/j.csi.2023.103802.

[16] P. Manjula and S. Baghavathi Priya, “An effective network intrusion detection and classification system for securing WSN using VGG-19 and hybrid deep neural network techniques,” *Journal of Intelligent and Fuzzy Systems*, vol. 43, no. 5, pp. 6419–6432, Sep. 2022, doi: 10.3233/JIFS-220444.

[17] G. Liu, H. Zhao, F. Fan, G. Liu, Q. Xu, and S. Nazir, “An enhanced intrusion detection model based on improved kNN in WSNs,” *Sensors*, vol. 22, no. 4, Feb. 2022, doi: 10.3390/s22041407.

[18] V. Sivagaminathan, M. Sharma, and S. K. Henge, “Intrusion detection systems for wireless sensor networks using computational intelligence techniques,” *Cybersecurity*, vol. 6, no. 1, Oct. 2023, doi: 10.1186/s42400-023-00161-0.

[19] S. Rameshkumar, R. Ganesan, and A. Merline, “Progressive transfer learning-based deep Q network for DDOS defence in WSN,” *Computer Systems Science and Engineering*, vol. 44, no. 3, pp. 2379–2394, 2023, doi: 10.32604/csse.2023.027910.

[20] V. Shakya, J. Choudhary, and D. P. Singh, “IRADA: integrated reinforcement learning and deep learning algorithm for attack detection in wireless sensor networks,” *Multimedia Tools and Applications*, vol. 83, no. 28, pp. 71559–71578, Feb. 2024, doi: 10.1007/s11042-024-18289-7.

[21] M. Al Lail, A. Garcia, and S. Olivo, “Machine learning for network intrusion detection—a comparative study,” *Future Internet*, vol. 15, no. 7, Jul. 2023, doi: 10.3390/fi15070243.

[22] M. Al-Shalabi, M. Anbar, T.-C. Wan, and A. Khasawneh, “Variants of the low-energy adaptive clustering hierarchy protocol: Survey, issues and challenges,” *Electronics*, vol. 7, no. 8, Aug. 2018, doi: 10.3390/electronics7080136.

[23] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 16–28, Jan. 2014, doi: 10.1016/j.compeleceng.2013.11.024.

[24] D. Asif, M. Bibi, M. S. Arif, and A. Mukheimer, “Enhancing heart disease prediction through ensemble learning techniques with hyperparameter optimization,” *Algorithms*, vol. 16, no. 6, Jun. 2023, doi: 10.3390/a16060308.

[25] L. R. Namamula and D. Chaytor, “Effective ensemble learning approach for large-scale medical data analytics,” *International Journal of System Assurance Engineering and Management*, vol. 15, no. 1, pp. 13–20, Feb. 2024, doi: 10.1007/s13198-021-01552-7.

BIOGRAPHIES OF AUTHORS

Mohamed Loughmari a Ph.D. student in computer science, technology and innovation, Engineering Sciences Laboratory (LSI), Polydisciplinary Faculty of Taza, Sidi Mohamed Ben Abdellah University, Taza/Morocco, received the master degree in intelligent and mobile systems from Sidi Mohammed Ben Abdellah University, Morocco, in 2021. His research interests include cybersecurity, artificial intelligence, machine learning, deep Learning, and computer networks. He can be contacted at email: mohamed.loughmari@usmba.ac.ma.

Anass El Affar is a professor of computer science at the Polydisciplinary Faculty of Taza, Sidi Mohamed Ben Abdellah University, Taza, Morocco. A member of the Engineering Sciences Laboratory (LSI), holds a Ph.D. in computer science. his research interests focus on computer science, intelligent system modeling, analysis and information processing, and artificial intelligence. He can be contacted at email: anass.elaffar@usmba.ac.ma.