

Enhancing intrusion detection systems with hybrid HHO-WOA optimization and gradient boosting machine classifier

Mosleh M. Abualhaj¹, Ahmad Adel Abu-Shareha², Roqia Rateb³

¹Department of Networks and Cybersecurity, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

²Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

³Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

Article Info

Article history:

Received Sep 26, 2024

Revised Apr 11, 2025

Accepted Jun 10, 2025

Keywords:

Feature selection

Gradient boosting machine

Harris Hawks algorithm

Intrusion detection system

Whale optimization algorithm

ABSTRACT

In this paper, we propose a hybrid intrusion detection system (IDS) that leverages Harris Hawks optimization (HHO) and whale optimization algorithm (WOA) for feature selection to enhance the detection of cyberattacks. The hybrid approach reduces the dimensionality of the NSL-KDD dataset, allowing the IDS to operate more efficiently. The reduced feature set is then classified using logistic regression (LR) and gradient boosting machine (GBM) classifiers. Performance evaluation demonstrates that the GBM-HHO/WOA combination outperforms the LR-HHO/WOA approach, achieving an accuracy of 97.68%. These results indicate that integrating HHO and WOA significantly improves the IDS's ability to identify intrusions while maintaining high computational efficiency. This research highlights the potential of advanced optimization techniques to strengthen network security against evolving threats.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Mosleh M. Abualhaj

Department of Networks and Cybersecurity, Faculty of Information Technology

Al-Ahliyya Amman University

Amman 19111, Jordan

Email: m.abualhaj@ammanu.edu.jo

1. INTRODUCTION

In recent years, as cyber threats have become more sophisticated and frequent, organizations have turned to intrusion detection systems (IDS) as a critical line of defense in protecting their networks and data. An IDS functions as a security tool that monitors and analyzes network traffic, identifying suspicious activity and potential threats in real-time [1], [2]. With the increasing volume of data generated by modern networks, traditional rule-based IDS approaches struggle to keep pace, often leading to higher false positives and negatives [2], [3]. As a result, machine learning (ML) techniques have emerged as a promising solution for enhancing IDS performance by enabling systems to learn from historical data and detect previously unknown attack patterns [1], [4].

The integration of ML classification algorithms into IDSs offers a dynamic approach to threat detection, allowing the system to differentiate between legitimate traffic and potential intrusions based on learned patterns [1], [4]. Classification algorithms such as gradient boosting machine (GBM) and logistic regression (LR) have shown strong potential in classifying network traffic data into benign or malicious categories [5], [6]. By leveraging the ability of ML models, IDSs can become more resilient and accurate in detecting a wide array of cyber threats, including both known and unknown attacks. However, the effectiveness of these models is highly dependent on the quality of the features used for training, which makes feature selection a critical component in building efficient IDSs [7], [8].

Feature selection algorithms play a key role in enhancing the performance of ML-based IDSs by identifying the most relevant and informative features from vast network traffic datasets. Reducing the dimensionality of the data not only improves model accuracy but also reduces computational costs, allowing for faster and more efficient intrusion detection. Optimization techniques such as whale optimization algorithm (WOA) and Harris Hawks optimization (HHO) have been employed to fine-tune feature selection processes, ensuring that the selected features maximize detection accuracy while minimizing noise and irrelevant data. The combination of effective feature selection and robust classification algorithms can significantly improve the reliability and performance of IDSs, offering a powerful tool for cybersecurity professionals to safeguard their networks against increasingly complex cyber threats [9]-[12].

Numerous research works have proposed to enhance IDS systems performance. Zhao and Zhao [13] proposed a solution that improves the accuracy of IDS systems by using ML techniques. The proposed solution uses the radial basis function (RBF) neural networks to extract important features from the data. Then, the support vector machine (SVM) technique is used for classification based on the extracted features from the RBF technique. The proposed solution was tested using the KDD99 dataset and implemented in Python. The results showed that combining RBF and SVM techniques achieved a high accuracy of 97%.

Research by Daoud *et al.* [14] emphasize the potential of ML to enhance the capabilities of IDS. The authors propose implementing various ML algorithms, specifically k-nearest neighbor (KNN), decision tree, and random forest, within the IDS framework. The goal is to measure their effectiveness in improving detection accuracy. The paper uses the K-Fold cross-validation method to enhance detection rates. The findings indicate that the random forest algorithm, particularly with 100 trees, achieved the highest accuracy of 92.65%, outperforming the other algorithms tested.

Akande *et al.* [15] present a novel hybrid algorithm that combines convolutional neural networks (CNN) and deep neural networks. This innovative approach aims to enhance the accuracy and effectiveness of intrusion detection, addressing the limitations of traditional methods in identifying network intrusions. This hybrid approach is designed to categorize network packets and identify intrusions, classifying them as either normal or malicious. The CNN achieved a high accuracy rate of 99.18%, outperforming other classifiers.

2. RESEARCH METHOD

2.1. NSL-KDD dataset

The NSL-KDD dataset will be used in this work to evaluate the proposed ML model. The NSL-KDD dataset used contains 148,517 attacks and benign records. In addition, the NSL-KDD dataset consists of 40 features. There are 38 different types of attacks categorized into ten types of DoS attack, six types of probe attack, seven types of users to root (U2R) attack, and 15 types of remote to local (R2L) attack. Table 1 list these 38 different types of attacks [16], [17].

Table 1. NSL-KDD dataset attack groups

Main attack	Subtypes
DoS	Back, Processtable, Pod, Land, Smurf, Neptune, Apache2, Teardrop, Udpstorm, and Worm
Probe	Mscan, Satan, Nmap, Ipsweep, Portsweep, and Saint
U2R	Loadmodule, Buffer overflow, Sqlattack, Xterm, Rootkit, Perl, and Ps
R2L	Guess_Password, Snpmpguess, Imap, Phf, Multihop, Warezmaster, Ftp_write, Xsnoop, Warez Spy, Sendmail, Xlock, Snpmpgetattack, client, Httptunnel, and Named

2.2. Logistic regression and gradient boosting machine classification algorithms

2.2.1. Logistic regression classifiers

LR is a binary classification algorithm that models the likelihood of an event occurring based on a collection of predictor factors using a logistic function. The logistic function generates a value between 0 and 1, which represents the likelihood that the event will occur [6], [18]. Figure 1 clarifies the LR technique.

2.2.2. Gradient boosting machine classifiers

GBM builds a sequence of DTs, each of which tries to correct the errors of the previous tree. This sequential process of building trees, adjusting the weights, and repeating the process is called boosting. The idea behind GBM is to combine multiple weak classifiers to create a strong classifier. A weak classifier is a simple model that performs only slightly better than random guessing. By combining many weak classifiers, GBM creates a powerful and accurate model [5], [19].

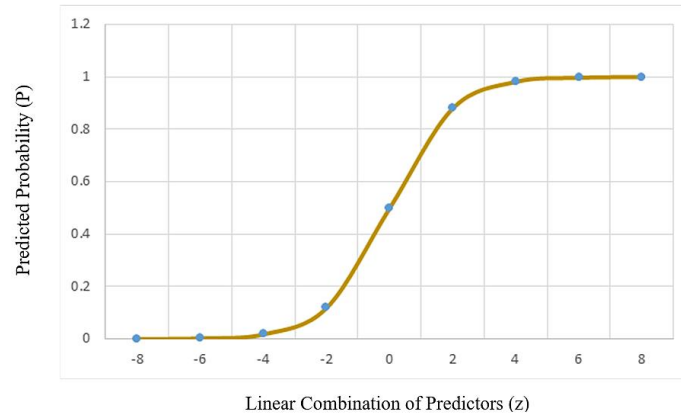


Figure 1. LR method

2.3. Harris Hawks optimization and whale optimization algorithm algorithms

The HHO and WOA algorithms have been used to select the key features of the attack from the NSL-KDD dataset. The HHO and WOA algorithms have been derived from the behavior of Harris Hawks and Whales, respectively [11], [12]. The HHO and WOA algorithms perform several operations to find the best solution. These operations are based on functions, variables, and constraints that impact finding the best solution. One of the key functions in HHO and WOA algorithms is the "objective function." The objective function for both HHO and WOA considers prediction accuracy for the feature subset and the number of selected features; this aims to choose a feature subset that enhances the model's predictive accuracy while avoiding overfitting by minimizing the number of features. The objective function of HHO and WOA is calculated using (1) [9], [10], [20]–[23].

$$\text{Objective function} = \text{Alpha} * \text{Error Rate} + \text{Beta} * \left(\frac{\text{number of selected features}}{\text{maximum number of features}} \right) \quad (1)$$

Alpha and Beta are weights that determine the relative importance of the error rate and the feature count. In this study, Alpha=0.99 and Beta=0.01, indicating a high emphasis on minimizing the error rate. The error rate is calculated as (1–Accuracy), where Accuracy is the performance measure of the classification model using the selected feature subset. This means that a lower error rate value indicated a better feature subset (better accuracy). The number of selected features is the count of features the algorithm chooses. The maximum number of features is the count of all available features in the dataset [9], [10], [20]–[23].

Besides the objective function, the decision variables form a key element of the HHO and WOA algorithms. In HHO and WOA, the decision variables are the positions of the hawks/whales in the multidimensional search space. Each hawk's/whale's position represents a potential solution (feature subset), where each dimension of this position represents a feature. This initially can be represented with a position vector [9], [10], [20]–[23], as (2) and (3):

$$\text{Vector } x_i = [x_{i,1}, x_{i,2}, x_{i,3} \dots x_{i,d}] \quad (2)$$

$$x_{i,d} = lb_d + (ub_d - lb_d) * \text{Rand}() \quad (3)$$

where $x_{i,d}$ is the position of i hawk in the d dimension, d is the number of features in the dataset, lb_d and ub_d are the lower and upper bounds for the d dimension (the lower bound is 0 and upper bound is 1), $\text{rand}()$ generates a random number between 0 and 1. The value indicates whether a feature is included in the subset being considered. Each position vector represents a feature subset (solution), and these vectors are updated using the exploitation and exploration phases of the algorithm. In the HHO exploration phase, the decision variables are the positions of the hawks, which are updated using (4) [9], [10], [20]–[23].

$$x_{i,d(\text{new})} = x_{i,d(\text{current})} + r * (x_{\text{random},d} + x_{i,d(\text{current})}) \quad (4)$$

Where $x_{i,d(\text{new})}$ is the updated position of the i hawk in the d dimension, $x_{i,d(\text{current})}$ is the current position of the hawk, r is a random number between 0 and 1, $x_{\text{random},d}$ is the position of a randomly selected hawk from the population. As for the exploitation phase, the positions of the hawks are updated based on the location of the prey (the best solution found so far) using (5) [9], [10], [20]–[23].

$$x_{i,d(new)} = x_{prey,d} + r * |x_{prey,d} + x_{i,d(current)}| \quad (5)$$

Where $x_{prey,d}$ is the current best solution found by any hawk in the population. In the WOA exploration phase, the decision variables are the positions of the whales, which are updated using (6) [9], [10], [20]–[23].

$$x_{i,d(new)} = x_{i,d(random)} - A * |C * x_{random,d} + x_{i,d(current)}| \quad (6)$$

Where $x_{i,d(new)}$ is the updated position of the i whale in the d dimension, $x_{i,d(current)}$ is the current position of the whale, $x_{random,d}$ is the position of a randomly selected hawk from the population, A and C are coefficients calculated as using the iterative decrement of parameter a . The exploitation phase of WOA includes two behaviors: encircling prey (7) and bubble-net attacking method (8) [9], [10], [20]–[23].

$$x_{i,d(new)} = x_{best,d} - A * |C * x_{best,d} + x_{i,d(current)}| \quad (7)$$

$$x_{i,d(new)} = |x_{best,d} - x_{i,d(current)}| * e^{b-l} * \cos(2\pi l) + x_{best,d} \quad (8)$$

Where b and l are constants and a random number, defining the shape of the spiral movement. In addition, the performance of HHO and WOA algorithms is impacted by some constraints. The primary constraint in both HHO and WOA is the boundary constraint, that the position of each hawk or whale stays within predefined lower and upper bounds for each dimension (feature), as shown in (9) [9], [10], [20]–[23].

$$x_{i,d} = \begin{cases} lb_d, & \text{if } x_{i,d} > lb_d \\ ub_d, & \text{if } x_{i,d} > ub_d \\ x_{i,d}, & \text{otherwise} \end{cases} \quad (9)$$

Another constraint is the binary conversion constraint, in which values are mapped to zero or one based on a threshold set to 0.5. When $x_{i,d} > \text{threshold}$, the feature should be included in the subset, so it's converted to 1. Otherwise, the feature shouldn't be included in the subset, so it's converted to 0. The features are represented in a binary vector. A position vector could be something like [0.2, 0.7, 0.6, 0.4, 0.8]. After applying binary conversion with a threshold of 0.5, this vector becomes [0, 1, 1, 0, 1]. The iteration limit and population size also impact HHO and WOA algorithms. The algorithm searches for better solutions as long as the current iteration number is less than the maximum number of iterations, and the population size limit the number of hawks/whales feature subsets that the algorithm can produce in each iteration [9], [10], [20]–[23].

2.4. The proposed intrusion detection system-based machine learning model

The proposed IDS model has gone through several processes to reach attack detection. First, the text data in the NSL-KDD dataset is converted into numbers using the label-encoding method. Then, the large number of data is narrowed to small ranges using the min-max scaler method [24], [25]. After that, the proposed hybrid features selection method is applied to select the most useful feature from the NSL-KDD dataset for attack detection. Finally, the LR and GBM classification algorithms are tested to classify the attacks. Figure 2 shows the proposed IDS model.

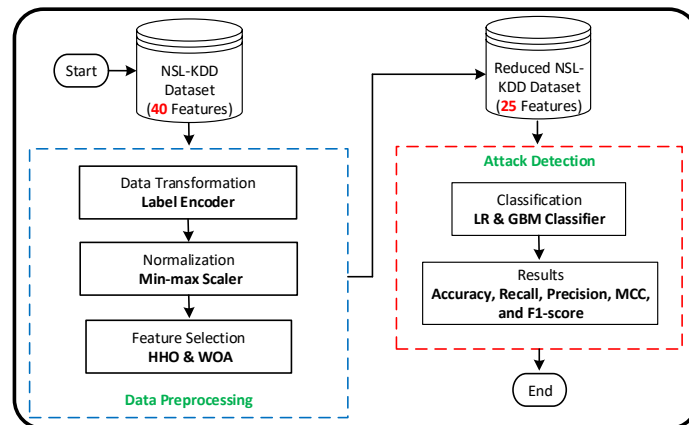


Figure 2. The proposed ML model

The proposed IDS model incorporates a hybrid feature selection approach by combining HHO and WOA to enhance the detection of network intrusions. Both algorithms are applied in parallel to the NSL-KDD dataset, selecting 13 features with HHO and 16 with WOA. By merging their common features, the feature space is reduced from 40 to 25, which improves the model's efficiency and reduces computational complexity. This hybrid method leverages the strengths of both algorithms, optimizing feature selection to improve the model's overall performance in detecting cyber threats. Table 2 shows the feature selected by HHO, WOA and HH/WOA methods.

Table 2. Selected feature by different optimizers

Optimizer	Selected features
WOA	Flag, src_bytes, num_root, num_outbound_cmds, is_host_login, Service, srv_count, num_failed_logins, serror_rate, dst_host_same_src_port_rate, num_access_files, srv_serror_rate, same_srv_rate, is_guest_login, srv_diff_host_rate, dst_host_rerror_rate
HHO	hot, Flag, protocol_type, src_bytes, urgent, dst_host_diff_srv_rate, num_access_files, diff_srv_rate, dst_host_count, dst_bytes, dst_host_srv_count, Count, dst_host_same_src_port_rate
HHO and WOA	src_bytes, dst_host_count, protocol_type, flag, dst_bytes, is_guest_login, urgent, hot, num_failed_logins, num_access_files, is_host_login, dst_host_same_src_port_rate, num_outbound_cmds, service, Count, serror_rate, srv_serror_rate, diff_srv_rate, srv_count, same_srv_rate, srv_diff_host_rate, dst_host_srv_count, dst_host_diff_srv_rate, num_root, dst_host_rerror_rate

3. RESULTS AND DISCUSSION

The confusion matrix (Figure 3) allows us to calculate various metrics that measure the performance of the classification model, including accuracy, recall, precision, matthew correlation coefficient (MCC), and F1-score. Accuracy can be defined as the proportion of correctly predicted attacks related to the total number of attacks that were forecasted. In (10) could be employed to determine the level of accuracy. The ratio of the number of samples in an attack class that can be accurately predicted to the total number of successful predictions for that attack class is referred to as the recall. In (11) can be used to determine recall of an event. The term precision refers to the ratio of the number of attacks that are accurately identified as attacks to the total number of attacks that are identified as attacks. In (12) can be employed to determine precision. F1-score is a metric that combines precision and recall to provide a single measure of a classification model's performance. It is the harmonic mean of precision and recall, calculated using (13) [24], [25].

$$Accuracy = \frac{(TPo+TNe)}{(TPo+TNe+FPo+FNe)} \quad (10)$$

$$Recall = \frac{TPo}{(TPo+FN)} \quad (11)$$

$$Precision = \frac{TPo}{(TPo+FPo)} \quad (12)$$

$$F1 - score = \frac{Precision \times Recall}{Precision + Recall} \quad (13)$$

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TPo)	False Negative (FNe)
Actual Negative	False Positive (FPo)	True Negative (TNe)

Figure 3. Confusion matrix

Figure 4 shows the accuracy accomplished by the suggested ML model. The LR-HHO technique has an accuracy of 91.01%, the LR-WOA technique has an accuracy of 88.79%, and the LR-HHO/WOA technique has an accuracy of 93.40%. The accuracy accomplished by the LR-HHO/WOA technique outperformed the accuracy accomplished by the LR-HHO technique and by the LR-WOA technique by 2.39% and 4.61%, respectively. Therefore, the suggested LR-HHO/WOA technique enhances the accuracy of the ML model attack detection. On the other hand, The GBM-HHO technique has an accuracy of 89.81%, the

GBM-WOA technique has an accuracy of 94.30%, and the GBM-HHO/WOA technique has an accuracy of 97.68%. The accuracy accomplished by the GBM-HHO/WOA technique surpassed the accuracy accomplished by the GBM-HHO technique and by the GBM-WOA technique by 7.87% and 3.38%, respectively. Therefore, the suggested GBM-HHO/WOA technique enhances the accuracy of the ML model attack detection. However, overall, the suggested GBM-HHO/WOA technique has outperformed the suggested LR-HHO/WOA technique in attack detection.

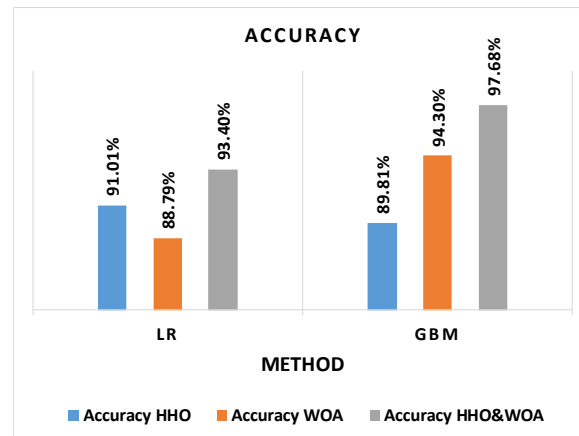


Figure 4. Accuracy of the proposed IDS-based ML model

Figure 5 shows the recall accomplished by the suggested ML model. The LR-HHO technique has a recall of 91.01%, the LR-WOA technique has a recall of 88.79%, and the LR-HHO/WOA technique has a recall of 93.40%. The recall accomplished by the LR-HHO/WOA technique outperformed the recall accomplished by the LR-HHO technique and by the LR-WOA technique by 2.40% and 4.62%, respectively. Therefore, the suggested LR-HHO/WOA technique enhances the recall of the ML model attack detection. On the other hand, The GBM-HHO technique has a recall of 98.81%, the GBM-WOA technique has a recall of 94.30%, and the GBM-HHO/WOA technique has a recall of 97.68%. The recall accomplished by the GBM-HHO technique surpassed the recall accomplished by the GBM-HHO/WOA technique and by the GBM-WOA technique by 1.13% and 4.51%, respectively. Therefore, the GBM-HHO technique enhances the recall of the ML model attack detection. However, overall, the suggested GBM-HHO technique has outperformed the suggested LR-HHO/WOA technique in attack detection recall accuracy.

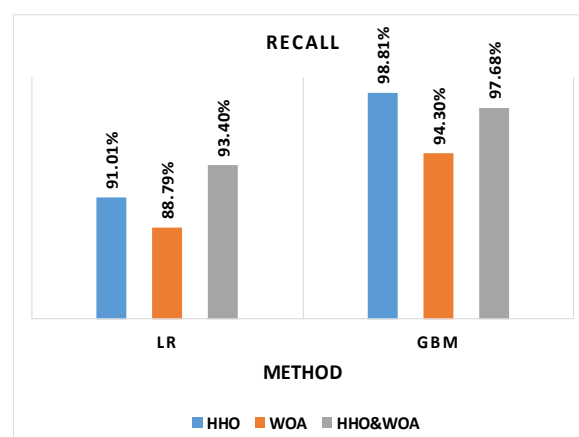


Figure 5. Recall of the proposed IDS-based ML model

Figure 6 shows the precision accomplished by the suggested ML model. The LR-HHO technique has a precision of 91.01%, the LR-WOA technique has a precision of 88.79%, and the LR-HHO/WOA technique has a precision of 93.40%. The precision accomplished by the LR-HHO/WOA technique

outperformed the precision accomplished by the LR-HHO technique and by the LR-WOA technique by 2.40% and 4.62%, respectively. Therefore, the suggested GBM-HHO/WOA technique enhances the precision of the ML model attack detection. On the other hand, The GBM-HHO technique has a precision of 89.81%, the GBM-WOA technique has a precision of 94.30%, and the GBM-HHO/WOA technique has a precision of 97.68%. The precision accomplished by the GBM-HHO/WOA technique surpassed the precision accomplished by the GBM-HHO technique and by the GBM-WOA technique by 7.87% and 3.38%, respectively. Therefore, the suggested GBM-HHO/WOA technique enhances the precision of the ML model attack detection. However, overall, the suggested GBM-HHO/WOA technique has outperformed the suggested LR-HHO/WOA technique in attack detection precision.

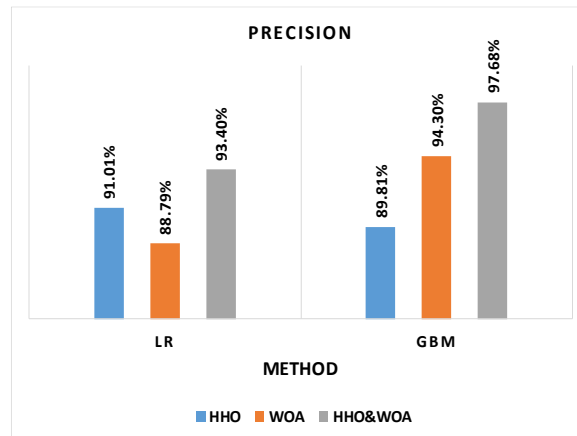


Figure 6. Precision of the proposed IDS-based ML model

Figure 7 shows the F1-score accomplished by the suggested ML model. The LR-HHO technique has an F1-score of 91.01%, the LR-WOA technique has an F1-score of 88.79%, and the LR-HHO/WOA technique has an F1-score of 93.40%. The F1-score accomplished by the LR-HHO/WOA technique outperformed the F1-score accomplished by the LR-HHO technique and by the LR-WOA technique by 2.40% and 4.62%, respectively. Therefore, the suggested GBM-HHO/WOA technique enhances the F1-score of the ML model attack detection. On the other hand, The GBM-HHO technique has an F1-score of 89.81%, the GBM-WOA technique has an F1-score of 94.30%, and the GBM-HHO/WOA technique has an F1-score of 97.68%. The F1-score accomplished by the GBM-HHO/WOA technique surpassed the F1-score accomplished by the GBM-HHO technique and by the GBM-WOA technique by 7.87% and 3.38%, respectively. Therefore, the suggested GBM-HHO/WOA technique enhances the F1-score of the ML model attack detection. However, overall, the suggested GBM-HHO/WOA technique has outperformed the suggested LR-HHO/WOA technique in attack detection F1-score.

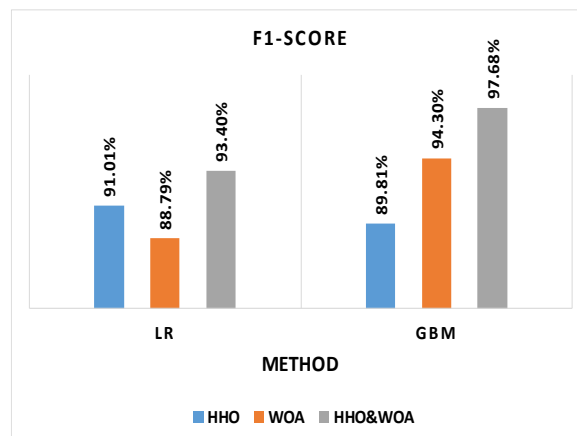


Figure 7. F1-score of the proposed IDS-based ML model

4. CONCLUSION

Cyberattacks are dangerous and pose a risk to individuals, corporations, and governments. There is a constant emergence of new methods for conducting cyberattacks. To counteract the effects of these cyberattacks, an IDS-based ML was suggested for this work. The IDS integrates HHO and WOA methods in order to lessen the size of the data that the IDS must handle. The performance of the suggested IDS model has been assessed using the NSL-KDD dataset with LR and GBM classifiers. The IDS model accomplished high performance when integrating HHO and WOA with LR and GBM classifiers. However, the suggested GBM-HHO/WOA technique has outperformed the suggested LR-HHO/WOA technique in attack detection recall accuracy. The accuracy accomplished by GBM-HHO/WOA is 97.68% while the accuracy achieved by LR-HHO/WOA is 93.40%. Overall, this research demonstrates the potential of HHO and WOA algorithms to improve the performance of IDSs significantly. By leveraging these advanced optimization techniques, the security and resilience of computer networks can be strengthened, ensuring effective protection against emerging cyber threats. Future works will compare the proposed model with other optimization feature selection methods, such as genetic algorithms or particle swarm optimization.




REFERENCES

- [1] B. Gao, B. Bu, W. Zhang, and X. Li, "An Intrusion Detection Method Based on Machine Learning and State Observer for Train-Ground Communication Systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6608–6620, Jul. 2022, doi: 10.1109/TITS.2021.3058553.
- [2] X. Y. Kong and G. H. Yang, "An Intrusion Detection Method Based on Self-Generated Coding Technology for Stealthy False Data Injection Attacks in Train-Ground Communication Systems," *IEEE Transactions on Industrial Electronics*, vol. 70, no. 8, pp. 8468–8476, Aug. 2023, doi: 10.1109/TIE.2022.3213899.
- [3] Y. Xue, J. Pan, Y. Geng, Z. Yang, M. Liu, and R. Deng, "Real-Time Intrusion Detection Based on Decision Fusion in Industrial Control Systems," *IEEE Transactions on Industrial Cyber-Physical Systems*, vol. 2, pp. 143–153, 2024, doi: 10.1109/ticps.2024.3406505.
- [4] R. Bitton and A. Shabtai, "A Machine Learning-Based Intrusion Detection System for Securing Remote Desktop Connections to Electronic Flight Bag Servers," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 3, pp. 1164–1181, May 2021, doi: 10.1109/TDSC.2019.2914035.
- [5] D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Gradient Boosting Feature Selection with Machine Learning Classifiers for Intrusion Detection on Power Grids," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 1104–1116, Mar. 2021, doi: 10.1109/TNSM.2020.3032618.
- [6] X. Wang, "Kronecker Factorization-Based Multinomial Logistic Regression for Hyperspectral Image Classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2021.3101509.
- [7] M. M. Abualhaj, A. A. Abu-Shareha, M. O. Hiari, Y. Alrabanah, M. Al-Zyoued, and M. A. Alsharaiah, "A Paradigm for DoS Attack Disclosure using Machine Learning Techniques," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 3, pp. 192–200, 2022, doi: 10.14569/IJACSA.2022.0130325.
- [8] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network Intrusion Detection for IoT Security Based on Learning Techniques," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 3, pp. 2671–2701, 2019, doi: 10.1109/COMST.2019.2896380.
- [9] M. Alazab, R. A. Khurma, P. A. Castillo, B. Abu-Salih, A. Martín, and D. Camacho, "An effective networks intrusion detection approach based on hybrid Harris Hawks and multi-layer perceptron," *Egyptian Informatics Journal*, vol. 25, no. 1, pp. 1–9, Mar. 2024, doi: 10.1016/j.eij.2023.100423.
- [10] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, May 2016, doi: 10.1016/j.advengsoft.2016.01.008.
- [11] M. M. Abualhaj, A. A. Abu-Shareha, A. Al-Allawee, A. Munther, and M. Anbar, "Performance Evaluation of Whale and Harris Hawks Optimization Algorithms with Intrusion Prevention Systems," in *International Conference on Soft Computing and Data Mining*, Cham: Springer Nature Switzerland, Putrajaya, Malaysia, Jul. 2024, pp. 254–265.
- [12] M. M. Abualhaj, S. N. Al-Khatib, A. Al-Allawee, A. Munther, and M. Anbar, "Enhancing Network Intrusion Detection Systems Through Dimensionality Reduction," in *International Conference on Soft Computing and Data Mining*, Cham: Springer Nature Switzerland, Putrajaya, Malaysia, Jul. 2024, pp. 244–253.
- [13] W. Zhao and Z. Zhao, "Providing a hybrid approach to increase the accuracy of intrusion detection systems in computer networks," *Journal of Engineering and Applied Science*, vol. 71, no. 1, pp. 1–19, Dec. 2024, doi: 10.1186/s44147-024-00404-y.
- [14] M. S. Daoud, B. A. Shawar, M. Alyafeai, A. Essam, M. Ghassan, and N. Alzaabi, "Enhancing Intrusion Detection Systems Accuracy Using Machine Learning," in *2023 10th International Conference on Software Defined Systems, SDS 2023*, IEEE, Oct. 2023, pp. 103–106, doi: 10.1109/SDS59856.2023.10329276.
- [15] H. B. Akande, C. Awoniyi, R. O. Ogundokun, A. A. Oloyede, O. A. Yiamiyu, and A. T. Caroline, "Enhancing Network Security: Intrusion Detection Systems with Hybridized CNN and DNN Algorithms," in *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)*, IEEE, Apr. 2024, pp. 1–7, doi: 10.1109/SEB4SDG60871.2024.10630078.
- [16] A. H. Mohammad, T. Alwada'n, O. Almomani, S. Smadi, and N. ElOmari, "Bio-inspired Hybrid Feature Selection Model for Intrusion Detection," *Computers, Materials and Continua*, vol. 73, no. 1, pp. 133–150, 2022, doi: 10.32604/cmc.2022.027475.
- [17] Ü. Çavuşoğlu, "A new hybrid approach for intrusion detection using machine learning methods," *Applied Intelligence*, vol. 49, no. 7, pp. 2735–2761, Jul. 2019, doi: 10.1007/s10489-018-01408-x.
- [18] F. Abramovich, V. Grinshtein, and T. Levy, "Multiclass classification by sparse multinomial logistic regression," *IEEE Transactions on Information Theory*, vol. 67, no. 7, pp. 4637–4646, Jul. 2021, doi: 10.1109/TIT.2021.3075137.
- [19] D. Rani, N. S. Gill, P. Gulia, F. Arena, and G. Pau, "Design of an Intrusion Detection Model for IoT-Enabled Smart Home," *IEEE Access*, vol. 11, pp. 52509–52526, 2023, doi: 10.1109/ACCESS.2023.3276863.
- [20] M. Z. Islam et al., "A Harris Hawks optimization based single and multi-objective optimal power flow considering environmental emission," *Sustainability*, vol. 12, no. 13, pp. 1–26, Jun. 2020, doi: 10.3390/su12135248.




- [21] O. Almomani, "A feature selection model for network intrusion detection system based on pso, gwo, ffa and ga algorithms," *Symmetry*, vol. 12, no. 6, pp. 1–20, Jun. 2020, doi: 10.3390/sym12061046.
- [22] M. Moukhafi, K. El Yassini, and S. Bri, "A novel hybrid GA and SVM with PSO feature selection for intrusion detection system," *International Journal of Advances in Scientific Research and Engineering*, vol. 4, no. 5, pp. 129–134, 2018, doi: 10.31695/ijasre.2018.32724.
- [23] O. Almomani, "A Hybrid Model Using Bio-Inspired Metaheuristic Algorithms for Network Intrusion Detection System," *Computers, Materials and Continua*, vol. 68, no. 1, pp. 409–429, 2021, doi: 10.32604/cmc.2021.016113.
- [24] A. Almomani *et al.*, "Ensemble-Based Approach for Efficient Intrusion Detection in Network Traffic," *Intelligent Automation and Soft Computing*, vol. 37, no. 2, pp. 2499–2517, 2023, doi: 10.32604/iasc.2023.039687.
- [25] Q. Y. Shambour, N. M. Turab, and O. Y. Adwan, "An Effective e-Commerce Recommender System Based on Trust and Semantic Information," *Cybernetics and Information Technologies*, vol. 21, no. 1, pp. 103–118, Mar. 2021, doi: 10.2478/cait-2021-0008.

BIOGRAPHIES OF AUTHORS






Prof. Mosleh M. Abualhaj    is a senior lecturer in Al-Ahliyya Amman University. He received his first degree in Computer Science from Philadelphia University, Jordan, in 2004, Master degree in Computer Information System from the Arab Academy for Banking and Financial Sciences, Jordan in 2007, and Ph.D. in Multimedia Networks Protocols from Universiti Sains Malaysia in 2011. His research area of interest includes VoIP, congestion control, cybersecurity data mining, and optimization. He can be contacted at email: m.abualhaj@ammanu.edu.jo.



Dr. Ahmad Adel Abu-Shareha    received his first degree in Computer Science from Al Al-Bayt University, Jordan, 2004, Master degree from Universiti Sains Malaysia (USM), Malaysia, 2006, and Ph.D. degree from USM, Malaysia, 2012. His research focuses on data mining, artificial intelligent, and multimedia security. He investigated many machine learning algorithms and employed artificial intelligent in variety of fields, such as network, medical information process, knowledge construction, and extraction. He can be contacted at email: a.abushareha@ammanu.edu.jo.



Dr. Roqia Rateb    has received her Ph.D. in Computer Science (artificial intelligence/data science) from University Utara Malaysia (UUM) in 2020. Her master has completed from Yarmouk University in Computer Information Systems (data mining) with Excellent. Moreover, she has received her Bachelor in Computer Science form Jordan University of Science and Technology with Excellent. She is currently an assistant professor at Department of Computer Science in Al-Ahliyya Amman University (AAU), Amman, Jordan. Her research interests are modelling dynamics of (multi) agent in practical application areas. She can be contacted at email: r.alshorman@ammanu.edu.jo.