SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I3P107

Volume 11 Issue 3, 86-96, March 2024

Original Article

A Novel Approach for Detection of DDoS Attacks in Software-Defined Networks Based on Grey Wolf Optimizer and Support Vector Machine

Aminata Dembele[1], Elijah Mwangi[2], Kennedy K. Ronoh[3], Edwin O. Ataro[4]

[1]*The Pan African University Institute for Basic Sciences, Technology and Innovation (PAUSTI), Kenya.*
[2]*Department of Electrical and Information Engineering, University of Nairobi, Kenya.*
[3]*School of Computing and Engineering Sciences, Strathmore University, Kenya.*
[4]*Department of Electrical and Information Engineering, Technical University, Kenya.*

[1]*Corresponding Author : aminata.dembele@students.jkuat.ac.ke*

Received: 05 January 2024 Revised: 04 February 2024 Accepted: 03 March 2024 Published: 25 March 2024

Abstract *- Software-defined networks face attacks that hinder efficient network provision and prevent users from accessing systems. Attack detection is crucial for better service provision and system resilience. Existing SDN-based Distributed Denial of Service (DDoS) detection technologies suffer from low accuracy, which is attributed to inadequate feature extraction and results in elevated false negative rates. This study introduces a solution leveraging the Grey Wolf Optimizer algorithm for feature selection to enhance DDoS attack detection and categorization. Employing a novel binary Grey Wolf optimization and Support Vector Machine (SVM) classifier on the InSDN dataset for SDNs, the proposed approach demonstrates superior performance, achieving 100% accuracy and recall. Feature selection with Binary Grey Wolf yields a 97% F1-Score using the unimodal equation and 100% accuracy, 96% recall, and a 98% F1-Score with the multi-modal equation, underscoring its efficacy in bolstering SDN security against DDoS attacks.*

Keywords *- Intrusion Detection System, SDN, DDoS attack, SVM, Grey Wolf Optimizer, Binary Grey Wolf Optimizer, InSDN dataset.*

1. Introduction

Software Defined Networking (SDN) is a mode of network management and control that involves the separation of the data plane, which is in charge of forwarding network traffic, from the control plane, which is in charge of network intelligence and decision-making [1]. The concept of segregating the control and data planes in SDN technology has significant advantages, including improved flexibility, cost efficiency, and streamlined administration.

However, it also brings new vulnerabilities [2]. These include Probe, Denial of Service (DoS), Distributed Denial of Service (DDoS), Root to Local Attacks (R2L), User to Root (U2R), etc. One frequent attack SDN is DDoS, which aims to block the utilization of a system, service, website, application, or any other network resource. Typically, the attack makes a system respond slowly or disable the system entirely. Attacks that originate from single sources are known as DoS attacks. Today, more common attacks are launched at a target from multiple sources but coordinated from a central point, known as DDoS attacks. Unlike DoS attacks, DDoS attacks are vast, potentially more devastating, and sometimes difficult for the victim to detect and stop. Figure 1 illustrates how an attacker performs a DDoS attack on a victim.

Monitoring network activity to find any malicious or unauthorized activity taking place on a computer network [4] is done using an Intrusion Detection System (IDS), which is a security tool. An IDS's primary function is to spot potential security lapses, such as hacking or vulnerabilities, and notify the network administrator or security staff so they may take the necessary action. This has led to the need to be able to detect such vulnerabilities at ease or at least improve on the current modes of spotting potential threats.

Recently, machine learning with enhanced feature selection has been utilized in Intrusion Identification Systems (IDS) to boost the identification of such threats [5]. The feature selection process chooses a subset of pertinent features (variables, attributes, or predictors) from a more extensive set of available features. Feature selection aims to find the most valuable and discriminative features significantly affecting a predictive model's performance or comprehension of the underlying data.

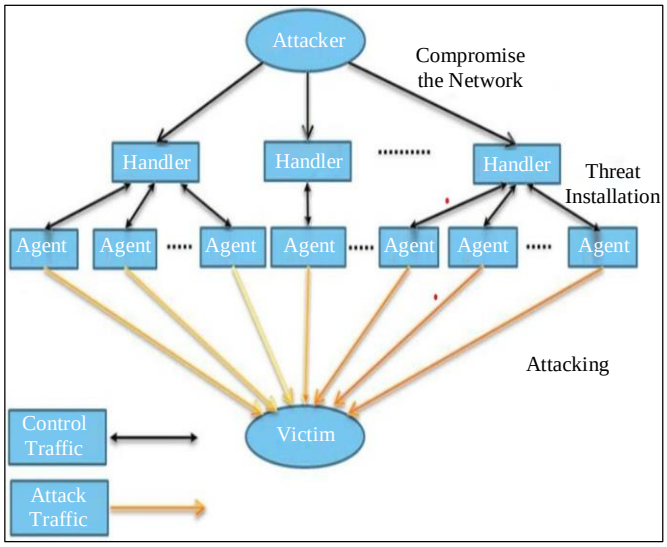

Fig. 1 DDoS attack illustration [3]

There are several ways of doing feature selection, and one technique is using the grey wolf optimizer; a metaheuristic algorithm is in a general area called metaheuristic algorithms. Network security is paramount in Software Defined Networks (SDNs) due to the increasing threat of Distributed Denial of Service (DDoS) attacks.

Despite the substantial efforts dedicated to this area, existing approaches exhibit limitations in effectively detecting and mitigating such attacks. This paper addresses a critical research gap by proposing a novel approach for accurate feature selection to detect DDoS attacks in SDNs. The current state of the art in DDoS attack detection within SDNs reveals a pressing need for innovative solutions.

Existing technologies for detecting DDoS attacks in SDN networks lack efficient feature extraction methods, resulting in low detection accuracy and high false-negative rates. Therefore, the metaheuristic algorithm must be used to improve feature selection. This study uses Grey Wolf Optimizer (GWO) and the Support Vector Machine (SVM) classifier to improve feature selection and the accuracy of DDoS attack detection and categorization. The experiment was conducted using the InSDN dataset for SDNs [14].

This research uses the InSDN dataset, a meticulously crafted dataset designed to comprehensively cover a broad spectrum of attacks across all SDN components within the proposed network testbed. Unlike other datasets, InSDN provides a unique and holistic representation of real-world threats in SDNs, ensuring the robustness and applicability of the proposed methodology.

When choosing the GWO [15] and SVM to detect Distributed Denial of Service (DDoS) assaults in Software Defined Networks (SDNs), we have carefully considered their distinct strengths and suitability for the situation at hand. The Grey Wolf Optimization (GWO) algorithm, a recently developed metaheuristic algorithm, demonstrates favourable characteristics such as a rapid convergence rate and effectiveness in optimizing intricate problems. Due to the ever-changing and progressive characteristics of DDoS attacks, the flexibility of GWO could potentially offer a benefit in detecting patterns and irregularities in network data.

SVM, however, is selected for its established superior performance in binary classification tasks and its capacity to manage high-dimensional data efficiently. SVM are highly effective in accurately distinguishing between different classes in convoluted feature spaces, which makes them well-suited for the complex nature of analyzing network traffic. Although decision trees and other machine learning models have shown effectiveness in specific applications, SVM's strong performance validates their choice for this particular task, especially in situations with non-linear bounds. In addition, the integration of GWO and SVM offers a unique and unexplored method for detecting DDoS attacks in SDNs. This highlights the originality and potential impact of this research.

The originality of this work is underscored by introducing a novel approach by employing binary GWO to enhance feature selection within SVM, thereby significantly improving the accuracy of DDoS attack detection and categorization in SDNs. Additionally, incorporating two functions, Unimodal and multi-modal equations, as fitness functions and using the InSDN dataset adds another layer of innovation to the research, contributing to the optimization process and further distinguishing this work in DDoS detection in SDNs.

To the best of our knowledge, the combination of GWO and SVM has not been applied for DDoS detection alongside using the InSDN dataset. The main contributions, hence, of this work:

(i) The introduction of an enhanced approach for identifying DDoS attacks in SDN controllers. This approach uses the Binary Grey Wolf Optimizer (BGWO) to select features to maximise classification accuracy.
(ii) Two equations, namely an unimodal equation and a multi-modal equation, were employed in the BGWO method to improve the feature selection performance and classification process.
(iii) To predict malicious traffic accurately, a model based on the SVM is employed to detect DDoS attacks in SDN.

This paper is arranged as follows. In Section 2, related work in the area of feature selection, as well as DDoS attacks, is presented. This is followed by the proposed model, discussed in Section 3. Section 4 discusses the experimental setup of this work. Section 5 presents and discusses the results. The work is concluded in Section 6.

2. Related Works

In recent years, various machine learning techniques have been explored to detect Distributed Denial of Service (DDoS) attacks in Software-Defined Networks (SDNs). This literature review aims to provide a more in-depth understanding of the existing approaches.

In [6], machine learning techniques used for feature selection of DDoS attacks have been reviewed. Neural networks, Naive Bayesian, Random Forest, KNN and SVM were evaluated, and Random forest performed better, having an accuracy of 98.70% with a weighted average of 98.4%. One of the main drawbacks of Random forest is that if it has many trees, the algorithm becomes too slow and ineffective for real-time predictions.

Also, Naive Bayes assumes that all predictors (or features) are independent; this rarely happens in real life. This limits the applicability of this algorithm in real-world use cases. This algorithm faces the zero-frequency problem where it assigns zero probability to a categorical variable whose category in the test data set was not available in the training dataset.

The study in [7] focuses on using machine learning methods to detect DDoS attacks to reduce misclassification errors. Using mutual information and random forest features, the authors found that random forest, gradient boosting, weighted voting ensemble, and KNN had better accuracy. Random Forest outperformed other classification methods, only misclassifying a single feature. However, KNN's main limitations include slow prediction stages, high memory requirements, and computational ineffectiveness due to its computational cost and extended training time.

In [8], an investigation was carried out to examine the effectiveness of RF, SVM, K-NN, and Naïve Bayes (NV) algorithms alongside Decision Tree (DT) algorithms. The evaluation of the NSL-KDD dataset demonstrated significant performance, with DT achieving an impressive accuracy of 99.97%. In contrast, SVM showed a notably lower accuracy of 60.19%. It is essential to highlight that the model proposed in this research underwent testing and training using artificial datasets that do not faithfully represent the unique characteristics of SDN networks.

The study in [9] proposes a Feature Selection (FS) technique using a Modified Binary Grey Wolf Optimizer (MBGWO) to improve Intrusion Detection (IDS) performance. The algorithm uses binary grey wolf optimization and an ideal number of features. The NSL-KDD network intrusion dataset was used to evaluate the technique.

The proposed FS and classification algorithms improved IDS performance, increasing intrusion detection accuracy to 99.22% and decreasing false alarm frequency. However, the algorithm requires intervention from all four wolves, potentially increasing the time to find the best solution.

In reference [10], authors examined various machine learning classification models such as DT, Random Forest (RF), AdaBoost (AB), Multilayer Perceptron (MLP), and Logistic Regression (LR) to analyze and detect TCP-SYN flood DDoS attacks on SDN controllers. The experimental results revealed that all the classification models exhibited outstanding performance. Nevertheless, it is crucial to emphasize that the proposed approach is explicitly tailored for addressing TCP-SYN flood attacks. The evaluation of this method was conducted using a relatively limited dataset.

In [11], a hybrid model for the SDN controller was introduced, which integrates an autoencoder and a one-class SVM to detect Distributed Denial of Service (DDoS) attacks. The model demonstrated an average accuracy of 99.35%. However, it comes with the drawback of introducing unnecessary load and overhead. Additionally, it is worth noting that the model was trained on an artificially created or generated dataset, potentially limiting its ability to represent the real-world SDN network environment accurately.

In [12], the hybrid GWO-PSO method presented in this research utilizes the NSL-KDD dataset for binary and multi-class challenges, showcasing the effectiveness of the proposed approach. The results show a remarkable accuracy of 99.97%, outperforming existing LSTM-RNN with its 97.72% accuracy. Additionally, the multi-class SVM achieved 98%, and the modified rank-based information gain feature selection method demonstrated an accuracy of 99.8%. However, it's crucial to acknowledge that the methodology is influenced by the system's complexity, indicating room for improvement to enhance overall performance and results.

The investigation in [13] suggested an efficient IDS for identifying probing attacks that utilized the Light Gradient Boosting Machine (LightGBM) and Grey-Wolf Optimizer (GWO) classifiers. The suggested IDS, deemed new, was trained and tested using the InSDN dataset; the suggested IDS assessment showed improved performance in probing attack detection within SDN compared to peer IDSs. Its performance was 99.8% for accuracy, 99.7% for recall, 99.99% for precision, and 99.8% for F-measure; the suggested IDS outperforms the most advanced IDSs.

Although they are based on machine learning, most of the studies mentioned depend on methods with a high rate of false alarms, requiring immediate management. One of the most effective approaches for addressing this is to perform good feature extraction before classification.

3. Algorithm Design

The following section discusses how the problem will be solved and the techniques employed.

3.1. Overview of Proposed Model for DDoS Detection Using GWO and SVM

This section considers a model using Binary Grey Wolf Optimization (BGWO) and SVM to detect the DDoS attack. The conceptual model for detecting the DDOS attack using SVM and GWO is illustrated in Figure 2. The first step is data cleaning, transforming, and normalizing the InSDN standard dataset.

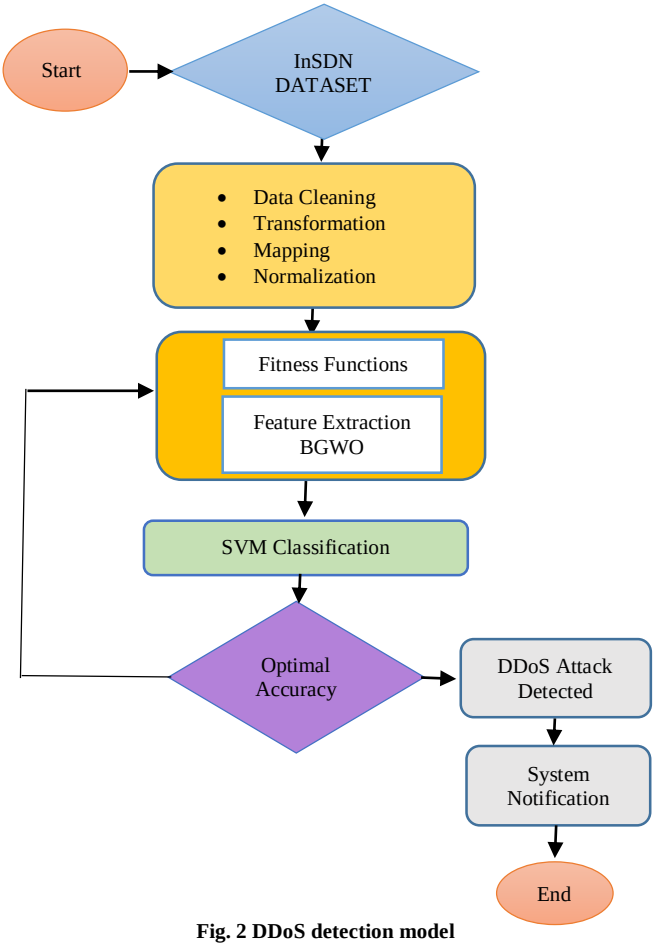

Fig. 2 DDoS detection model

The InSDN [14] public access dataset has been used in this study, where the data was cleaned by removing blank spaces, incorrect data entries, duplicates, etc from the dataset. This was followed by normalization, which was given by Equation 1:

$$X_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Where x represents the element in the dataset, X_{norm} is the normalized element, $x_{min,}$ and x_{max} are the minimum and maximum data, respectively. There was also a change in the label of the attacks to 1 and 0, where 1 is the DDOS attack, and 0 is every other attack. The second step is applying BGWO for feature selection. The final step is to classify the attack using SVM.

3.2. Grey Wolf Optimizer

GWO [15] is an algorithm that borrows from real-life grey wolf hunting tactics in Equation 5, which involves encircling, hunting, and attacking the prey. The steps are explained below.

3.2.1. Encircling the Prey

Grey wolves surround their prey in a circular formation while hunting. The following equations are presented to model encircling behaviour mathematically:

$$\vec{D} = \left|\vec{C}.\overrightarrow{X_p} - \vec{X}(t)\right| \quad (2)$$

and

$$\vec{X}(t+1) = \overrightarrow{X_p}(t) - \vec{A}.\vec{D} \quad (3)$$

Where t is the number of iterations, $\vec{A}$ and $\vec{C}$ are coefficient vectors, $\vec{X}_P$ is the vector position of the prey, $\vec{X}$ is the vector position of the grey wolf. Then, the coefficient vectors $\vec{A}$ and $\vec{C}$ can be expressed as:

$$\vec{A} = 2.\vec{a}\overrightarrow{r_1} - \vec{a} \quad (4)$$

$$\vec{C} = 2.\overrightarrow{r_2} \quad (5)$$

$\vec{a}$ is a set of vectors reduced linearly from 2 and 0, and r_1 and r_2 are randomly generated vectors within the range of 0 to 1.

3.2.2. Hunting the Prey

The distance between a particular wolf and the prey is given below.

$$\overrightarrow{D_\alpha} = \left|\vec{C}_1\overrightarrow{X_\alpha} - \vec{X}\right| \quad (6)$$

$$\overrightarrow{D_\beta} = \left|\vec{C}_1\overrightarrow{X_\beta} - \vec{X}\right| \quad (7)$$

$$\overrightarrow{D_\delta} = \left|\vec{C}_1\overrightarrow{X_\delta} - \vec{X}\right| \quad (8)$$

Where α represents the group leader in the pack, β is the second-best, and δ is the third-best leader. The vector positions of the grey wolves are given below:

$$\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \overrightarrow{A_1}\overrightarrow{D_\alpha} \quad (9)$$

$$\overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{A_1}\overrightarrow{D_\beta} \quad (10)$$

$$\overrightarrow{X_3} = \overrightarrow{X_\delta} - \overrightarrow{A_1}\overrightarrow{D_\delta} \quad (11)$$

3.2.3. Attacking the Prey

The following expression gives the attack on the prey:

$$\vec{a} = 2 - t\left(\frac{2}{maximum\ number\ of\ iterations}\right) \quad (12)$$

Where t represents the number of iterations, updating the wolf location is done using the following equation:

$$X(Iteration) = \frac{X_1+X_2+X_3}{3} \quad (13)$$

The GWO steps are given in Algorithm 1 below:

Algorithm 1: Grey Wolf Optimizer

Initialize the grey wolf population. X_i(i=1,2,…,n)
Initialize $\vec{a}$, A and C
Compute the fitness of each wolf
Set $\vec{X}_\alpha$ as the best wolf
Set $\vec{X}_\beta$ as the second-best wolf.
Set $\vec{X}_\delta$ as the third-best wolf.
while (t < $MaxIter$ number of iterations)
for each wolf, do
Using Equation 13, update the position of the current search agent
end for
Update a, A and C
Compute the fitness of all search agents
Update $\vec{X}_\alpha$, $\vec{X}_\beta$and $\vec{X}_\delta$
$t = t + 1$
end while
return $\vec{X}_\alpha$

3.3. Binary Grey Wolf for Feature Selection

According to [12], GWO can be modified to BGWO using binary operators such as the sigmoid, crossover or tanh functions for feature selection. The BGWO uses the three wolves, α, β, and δ. In BGWO, the position of the wolf is given by:

$$\omega_i^d = \begin{cases} 1 & if\ \left(\omega_i^d + step\ b_j^d\right) \geq 1 \\ 0 & Otherwise \end{cases} \quad (14)$$

$$step\ b_j^d = \begin{cases} 1 & if\ step\ C_j^d \geq random \\ 0 & Otherwise \end{cases} \quad (15)$$

$$step\ b_j^d = \frac{1}{1+e^{-10\left(A_i^d D_j^d - 0.5\right)}} \quad (16)$$

Where $i \in \{1,2,3\}$ and $j \in \{\alpha,\beta,\delta\}$ such that whenever, $i = 1$, then $j = \alpha$ and when $i = 2$, then $j = \beta$ and finally when $i = 3$, then $j = \delta, wd$ is then given by:

$$\omega_d = \begin{cases} w_1^d & if\ rand < \frac{1}{3} \\ w_2^d & if\ \frac{1}{3} \leq rand < \frac{2}{3} \\ w_3^d & Otherwise \end{cases} \quad (17)$$

Algorithm of BGWO is given below:

Algorithm 2: Binary Grey Wolf Optimizer

Initialize the grey wolf population. X_i(i=1,2,…,n)
Initialize $\vec{a}$, A and C
Evaluate the fitness of each wolf using Equations 18, and 19.
Set $\vec{X}_\alpha$ as the best wolf
Set $\vec{X}_\beta$ as the second-best wolf.
Set $\vec{X}_\delta$ as the third-best wolf.
while (t < $MaxIter$ number of iterations)
for i=1 to the number of wolves, n
Compute $\overrightarrow{X_1}$, $\overrightarrow{X_2}$, $\overrightarrow{X_3}$ using Equations 9, 10, and 11
Generate X_i^{new} by applying the Equation 15
next i
Evaluate the fitness of all grey wolves,
Update $\vec{X}_\alpha$, $\vec{X}_\beta$and $\vec{X}_\delta$
Update a, A and C
$t = t + 1$
end while
return $\vec{X}_\alpha$

The first fitness function, Equation 18, to be used in this research is a uni-modal, and the second function, Equation 19, is a multi-modal function from the benchmark functions used by [7].

$$f(x) = \sum_{i=1}^{n} x_i^2 \quad (18)$$

$$f(x) = \frac{1}{4000} + \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \quad (19)$$

Where x represents the positions and x_i represents the current position.

There are n features in any given dataset, implying that there are 2^n combinations for a given classification. The aim of feature selection before classification is to minimize the number of features used for classification to find a subset of feature combinations that could be used to maximize the classification accuracy [7].

3.4. Support Vector Machine (SVM)

SVM is a classifier that can detect attacks [16]. According to [17], SVM is the best-known technique used in data classification and regression to optimize the expected solution. It is usually used in solving binary classification problems. It typically minimizes the classification errors of the training data to obtain a better generalization ability. The SVM algorithm creates a hyperplane separating the data into two classes.

This is followed by finding points closest to the plane from both classes. The identified points will be support vectors. Once the support vectors are identified, the margin (distance between the plane and the support vectors) is maximized.

Suppose the training data set M = $\{(x_1, y_1), \cdots, (x_n, y_n)\}$ where $x_i \in R^n$ is the input vector and and y = {0, 1} is the target vector. Algorithm 3 illustrates how SVM works:

Algorithm 3: SVM

Require: x and y are loaded for training with the labeled dataset, $\alpha = 0$ or $\alpha = partially$ trained SVM.
C = some values (20 for example)
for { x_i, y_i } , { x_j, y_j } do
Optimize α_i and α_j
end for
Until no changes in α or other resource constraint criteria are met
Ensure: Retain only the support vectors ($\alpha_i > 0$)

3.5. Proposed BGWO-SVM Approach

3.5.1. A Brief Textual Description of the Proposed BGWO-SVM Approach and Its Overall Algorithm

The proposed BGWO-SVM approach combines the GWO and SVM to create an innovative algorithm for DDoS attack detection in SDNs.

1. The algorithm begins with initialising a population of grey wolves, representing potential solutions in the search space. These wolves undergo an iterative optimization process guided by unimodal and multi-modal equations serving as fitness functions, promoting exploring diverse solutions. As the optimization progresses, the fittest solutions are selected based on their performance in minimizing the chosen fitness function.
2. Subsequently, the optimized features are fed into the SVM classifier for the final classification task. SVM leverages the selected features to effectively distinguish between normal and malicious network traffic. The entire process is iteratively executed, with the Grey Wolf Optimizer continually refining the feature set to enhance the overall accuracy of the SVM-based DDoS detection.

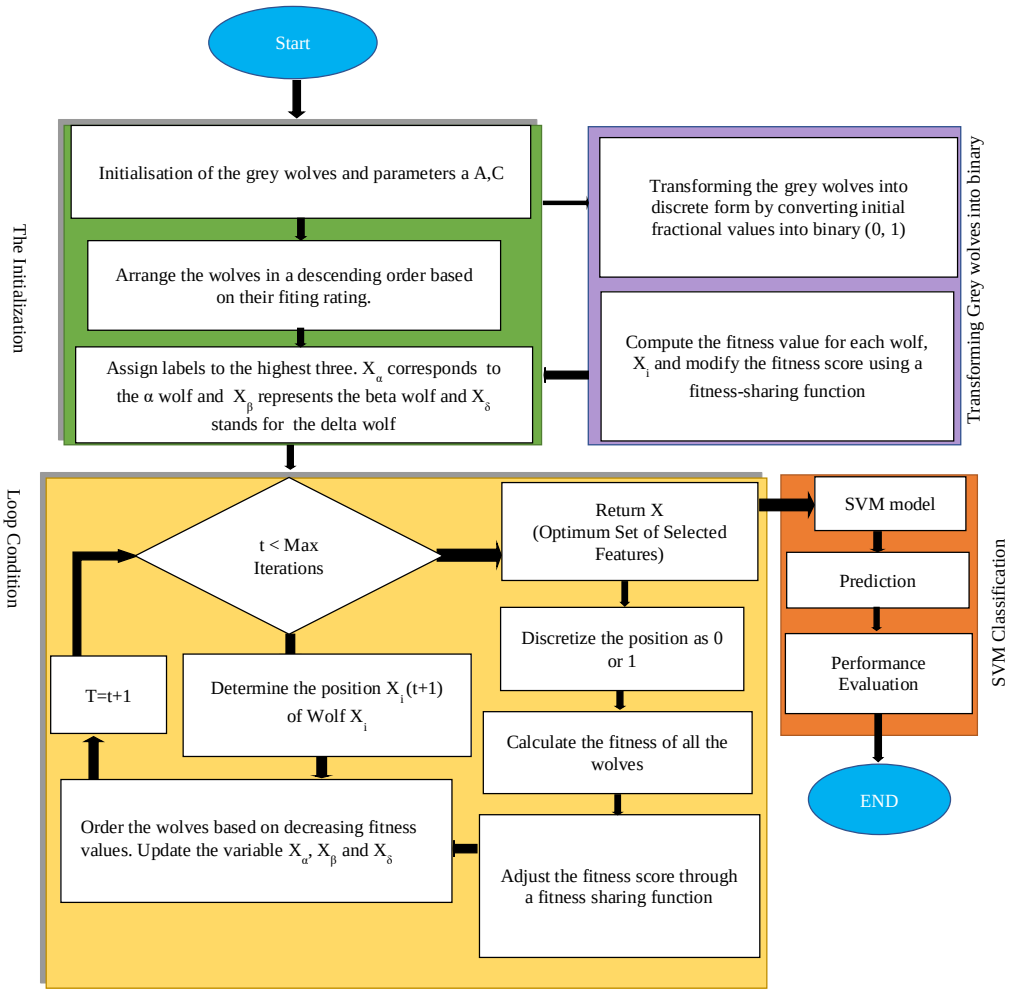

Fig. 3 Flowchart of DDoS detection model

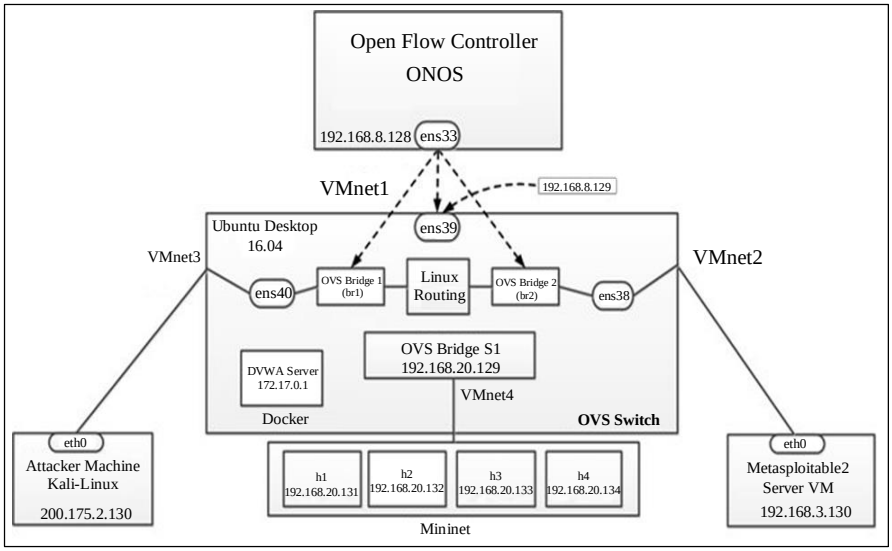

Fig. 4 Logical network topology [13]

This innovative integration of metaheuristic optimization and machine learning classification aims to provide a robust and adaptive solution for detecting DDoS attacks in SDNs, as reflected in Figure 3, which shows the overall algorithm.

4. Experimental Setup

This section discusses the experiment setup of this work.

4.1. Dataset

The data set used in this study is the InSDN dataset [14], which is a standard data set for SDN attacks and was generated using Mininet simulation. The dataset contains various attacks which can occur in an SDN environment.

The total number of features in the dataset was 67. The new IDS was coded using Python programming language. The experiments were conducted on a personal computer with the following specifications: 8GB of RAM, Windows 11 64-bit, and a 1.6 GHz Intel Core (TM) i5-8th Generation processor. Table 1 shows the Mininet simulator configuration parameters.

Table 1. InSDN mininet setup configuration [13]

Mininet and OVS Switch Parameters Configuration	
Hosts Interface	Four Virtual Hosts (h1 to h4)
Remote Controller	Four Adapters in the OVS-VM,ens38 to ens41. Open Flow Controller ONOS.
Protocols	UDP, TCP and ICMP.
Switch	Default OVS Switch.
Link Adjustment	Connect the Kali Linux VM with the Same Adapter of br1 and Metasploitable2server with the br2 Adapter.

5. Performance Evaluation

This section discusses the Performance Matrix, Performance Measures, and the Results and Analysis

5.1. Performance Matrix

The evaluation matrix was obtained by testing the following parameters.

5.1.1. Accuracy

Accuracy is the main and most basic performance measure, which is the proportion of the correctly predicted observations to all observations. The accuracy formula is given in Equation 20.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (20)$$

Where,
TP is True Positive,
TN is True Negative,
FP is False Positive, and
FN is False Negative.

According to [18], the optimal accuracy should be greater than 99%.

5.1.2. Precision

Precision is the ratio of positive, accurately predicted observations to all positive expected observations. High precision is linked with a low false positive rate. The outcome of precision is the probability of how the classifier is predicting the positive class. The precision is given as in Equation 21.

$$Precision = \frac{TP}{TP+FP} \quad (21)$$

5.1.3. Recall

This is the ratio of accurately predicted positive observations in the actual class. It is computed as per Equation 22.

$$Recall = \frac{TP}{TP+FN} \quad (22)$$

5.1.4. F− Measure or F1 Score

It is a normalized average of precision and recall. This implies that the score includes both false positives and false negatives. Although it is more straightforward than accuracy, it's more valuable, mainly if the class distribution is irregular. It can be computed using Equation 23.

$$F1\ Score\ = 2\left(\frac{Precision \times Recall}{Precision+Recall}\right) \quad (23)$$

The values used for the evaluation matrix evaluations are obtained from the confusion matrix in Table 2.

Table 2. Confusion matrix

Confusion Matrix		
Scenarios	DDoS Attack	Not a DDoS Attack
DDoS Attack	TP	FP
Not a DDoS Attack	FN	TN

The results obtained from detecting the DDoS attack with or without feature selection are discussed below when considering the evaluation matrix in the previous section. The selected features used in the classification when feature selection was considered were 30 features, which is 45.45% of the total features. On the other hand, all the features were considered when feature selection was not used.

5.2. Performance Measures

The two scenarios were considered; with or without feature selection, when using SVM for classification. The performance with feature selection was 99.86%, and without feature selection was 98.42% when the uni-modal Equation 18 was used as the fitness function.

When the multi-modal Equation 19 was used as the fitness function, the performance with and without feature selection was 99.89% and 98.42%, respectively. From the two cases, it is clear that with feature selection, the algorithm performed better, and the multi-modal equation also performed better than the uni-modal one.

5.2.1. Confusion Matrix with and without Feature Selection

The confusion matrix of an unimodal and the multi-modal equations' results are presented as follows in Table 3.

5.2.2. Classification Report of SVM

The classification report obtained from the confusion matrix is shown in Table 4.

Table 3. Confusion matrix summary

Confusion Matrix				
	After Feature Selection		**Before Feature Selection**	
Equation Type	**Unimodal**	**Multimodal**	**Unimodal**	**Multimodal**
TP	15784	15805	15807	15807
FP	23	2	0	0
FN	0	16	255	255
TN	372	356	117	117

Table 4. Classification report of SVM with and without FS

Clasification Report				
	After Feature Selection		**Before Feature Selection**	
Equation Type	**Unimodal**	**Multimodal**	**Unimodal**	**Multimodal**
Accuracy	100%	100%	98%	98%
Precision	94%	99%	100%	100%
Recall	100%	96%	31%	31%
F1-Score	97%	98%	48%	48%

5.3. Results and Analysis

The following classification report was obtained from the confusion matrix. For the unimodal Equation 18, it can be noted that there is better performance with feature selection when it comes to accuracy, where it is 100% with FS and 98% without FS. Recall is 100% with FS and 31% without FS, and F1-Score is 97% with FS and 48% without FS when detecting DDoS attacks. The only parameter that performed worse than the one with feature selection is precision, with 94% with feature selection and 100% without feature selection. The classification report of SVM with FS without FS is shown in Table 4, and the information is presented in Figure 5.

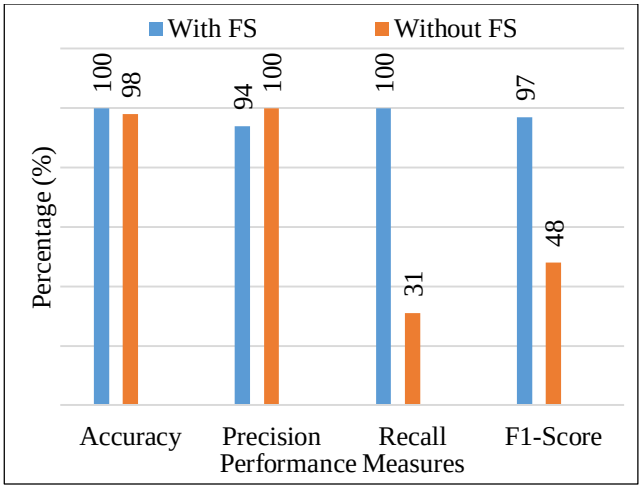

Fig. 5 Classification report for DDos attack using equation 18 as the fitness function

On the other hand, the Multi-modal Equation 19 shows better performance with feature selection regarding accuracy, where it is 100% with FS and 98% without FS. While recall is 96% with FS and 31% without FS, and F1-Score is 98% with FS and 48% without FS when detecting DDoS attacks. The only parameter that performed worse than the one with feature selection is precision, with 99% with feature selection and 100% without feature selection. The performance of SVM with respect to an unimodal function and with a multi-modal function, as seen in Table 4, shows that The accuracy in both is better than without feature selection.

(i) The accuracy in both is better than without feature selection
(ii) Precision in the multi-modal equation is better than in the unimodal equation; however, they perform worse than without feature selection.
(iii) Recall, on the other hand, is better with the unimodal equation with feature selection.
(iv) F1-Score with the multi-modal equation is better with feature selection.

The classification report for DDoS attack is presented in Figure 6.

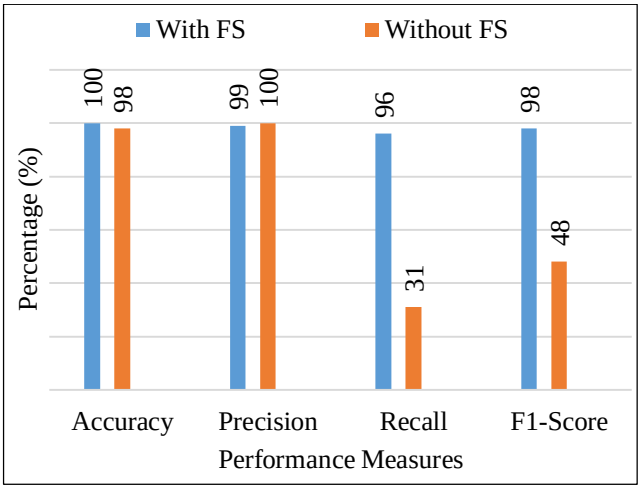

Fig. 6 Classification report for DDos Attack using equation 19

5.4. Comparative Analysis

Table 5 compares the results produced with our approach versus existing methodologies. Our proposed method was compared to previous works to evaluate its effectiveness. The proposed BGWO-SVM obtained a better accuracy of 100% compared to existing methods in the literature review, such as [6-13].

Based on the comparison of approaches employing the same InSDN dataset, the suggested IDS outperforms all current IDSs. Figure 7 below shows a comparison of performance measurements.

Table 5. Comparative analysis between our new approach and existing method

References	Methodologies	Dataset	Accuracy (%)
[6]	Neural networks, NB, RF, KNN and SVM	**InSDN**	98.70
[7]	Mutual Information and Random Forest, KNN	CICIDS 2017- CICDDoS 2019	99.99
[8]	RF, SVM, K-NN, and NV , DT	NSL-KDD	99.97
[9]	MBGWO	NSL-KDD	99.22
[10]	Ml based adaboost	created dataset	99.99
[11]	SAE-1SVM	CICIDS20127 dataset	99.35
[12]	Hybrid GWO-PSO	NSL-KDD	99.98
[13]	BGWO-LightGBM	**InSDN**	99.8
Proposed Method	BGWO-SVM	**InSDN**	100

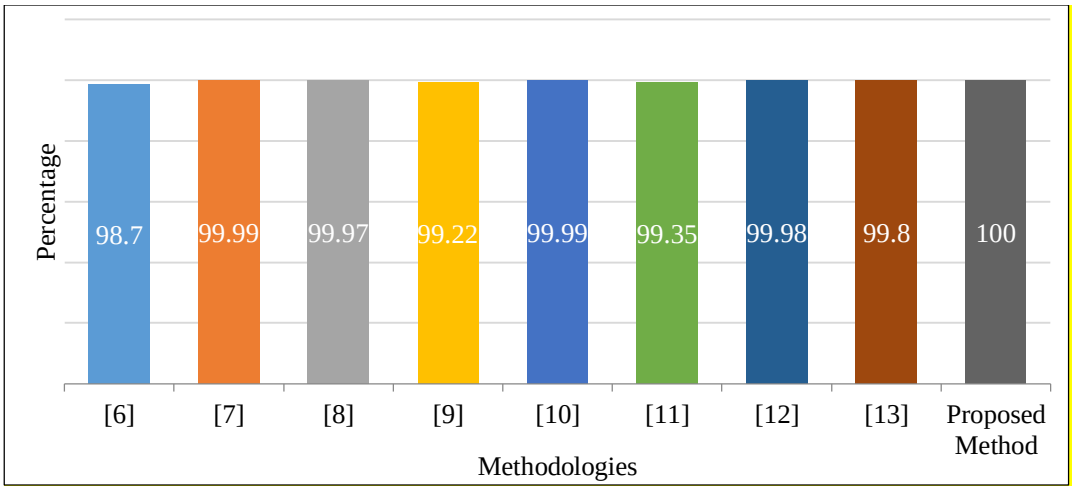

Fig. 7 Comparison of performance measurements between existing and proposed methods

6. Conclusion

This study has shown that when using only 45.45% and 50% of the total 66 features, the uni-modal and multi-modal functions perform better than when using all the features of the corresponding dataset. The algorithm performs better with the improved feature selection than the one without feature selection. The accuracy, recall and F1-Measure performance were also way better with the feature selection than without it. The precision measure did not perform better with FS compared to without FS, which was 94% for the feature selection algorithm with unimodal equation 99% for the feature selection algorithm with multi-modal equation and 100% without FS. In future research, we plan to apply the hybrid technique to improve accuracy and use fewer features for classification.

Acknowledgments

I acknowledge the invaluable contribution of Mawazo Institute and its partners in supporting my research, which has been instrumental in enabling me to achieve this project.

References

[1] Fetia Bannour, Sami Souihi, and Abdelhamid Mellouk, "Distributed SDN Control: Survey, Taxonomy, and Challenges," *Communications Surveys and Tutorials*, vol. 20, no. 1, pp. 333-354, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ali Nadim Alhaj, and Nitul Dutta, "Analysis of Security Attacks in SDN Network: A Comprehensive Survey," *Contemporary Issues in Communication, Cloud and Big Data Analytics*, pp. 27-37, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Wesam Bhaya, and Mehdi Ebady Manaa, "Review Clustering Mechanisms of Distributed Denial of Service Attacks," *Journal of Computer Science*, vol. 10, no. 10, pp. 2037-2046, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[4] Ansam Khraisat et al., "Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges," *Cybersecurity*, vol. 2, no. 1, pp. 1-22, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] Hongyu Liu, and Bo Lang, "Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey," *Applied Sciences*, vol. 9, no. 20, pp. 1-28, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] Andi Maslan et al., "Feature Selection for DDoS Detection Using Classification Machine Learning Techniques," *IAES International Journal of Artificial Intelligence*, vol. 9, no. 1, pp. 137-145, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Mona Alduailij et al., "Machine-Learning-Based DDoS Attack Detection Using Mutual Information and Random Forest Feature Importance Method," *Symmetry*, vol. 14, no. 6, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mahmoud Said ElSayed et al., "A Novel Hybrid Model for Intrusion Detection Systems in SDNs Based on CNN and A New Regularization Technique," *Journal of Network and Computer Applications*, vol. 191, pp. 1-18, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Qusay M. Alzubi et al., "Intrusion Detection System Based on a Modified Binary Grey Wolf Optimisation," *Neural Computing and Applications*, vol. 32, pp. 6125-6137, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Rochak Swami, Mayank Dave, and Virender Ranga, "Detection and Analysis of TCP-SYN DDoS Attack in Software-Defined Networking," *Wireless Personal Communications*, vol. 118, pp. 2295-2317, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Lotfi Mhamd et al., "A Deep Learning Approach Combining Autoencoder with One-Class SVM for DDoS Attack Detection in SDNs," *2020 IEEE Eighth International Conference on Communications and Networking (ComNet)*, Hammamet, Tunisia, pp. 1-6, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Ediga Sathyanarayana Phalguna Krishna, and Thangavelu Arunkumar, "Hybrid Particle Swarm and Gray Wolf Optimization Algorithm for IoT Intrusion Detection System," *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 4, pp. 66-76, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Abdulaziz Almazyad, Laila Halman, and Alaa Alsaeed, "Probe Attack Detection Using An Improved Intrusion Detection System," *Computers, Materials & Continua*, vol. 74, no. 3, pp. 4769-4784, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Mahmoud Said Elsayed, Nhien-An Le-Khac, and Anca D. Jurcut, "InSDN: A Novel SDN Intrusion Dataset," *IEEE Access*, vol. 8, pp. 165263-165284, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Andrew Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[16] Qasem Al-Tashi et al., "A Review of Grey Wolf Optimizer-Based Feature Selection Methods for Classification," *Evolutionary Machine Learning Techniques*, pp. 273-286, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] E. Emary, Hossam M. Zawbaa, and Aboul Ella Hassanien, "Binary Grey Wolf Optimization Approaches for Feature Selection," *Neurocomputing*, vol. 172, pp. 371-381, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[18] P. Karthika, and Karmel Arockiasamy, "Simulation of SDN Mininet and Detection of DDoS Attack Using Machine Learning," *Bulletin of Electrical Engineering and Information*, vol. 12, no. 3, pp. 1797-1805, 2023. [CrossRef] [Google Scholar] [Publisher Link]