Optimized Intrusion Detection for Cloud Networks: A Genetic Algorithm-Driven Approach to DDoS Mitigation

Reddy A. Hariprasad*, Professor, Dept.of CSE,Geethanjali College of Engineering and Technology.
K.Pushpa Rani**, Associate Professor, Department of CSE, MLR Institute of Technology.
Somala Rama Kishore, Associate Professor,Dept.of ECE, CMR Engineering College.

Abstract: As computing capabilities have expanded and cloud services have become more prevalent, the frequency and intensity of cyber-attacks, particularly Distributed Denial of Service (DDoS) attacks, have escalated. These attacks, which can target multiple hosts simultaneously, pose a significant challenge due to the decentralized nature of cloud-based service models. This study introduces an innovative Intrusion Detection System (IDS) designed to enhance network performance by accurately identifying DDoS attacks within cloud environments. The proposed IDS framework incorporates a genetic algorithm-based arithmetic optimization technique for selecting attack vectors and utilizes an intelligent ensemble Voting Classifier built on the Woodpecker-based Flamingo Search optimization algorithm. Preliminary results indicate that the proposed IDS significantly surpasses existing methods in terms of network performance, DDoS attack prediction, and mitigation effectiveness.

Keywords: DDOS attacks, Cloud environment, Intrusion detection system, Optimization, Genetic algorithm

1. Introduction:

Using cloud computing as a service version has won big recognition in current years. It makes it easy to access a ramification of shared computing sources, which includes garage, processing electricity, and programs, via the internet at any time and from everywhere within the globe. This accessibility is furnished around the clock. Those sources can be assigned to users on call for and made available to them in no time in any respect, and they handiest take a minimum amount of labor to keep [1]. It allows the corporations to dynamically adjust their abilties in a fee-powerful manner, which frees them from the weight of having to maintain an luxurious facts era infrastructure. Focusing on business enterprise's efforts on what it does first-rate—its center business—can assist that business enterprise acquire higher ranges of productivity. Inside the maximum latest few years, cloud computing has visible a meteoric upward push in recognition. A take a look at report that was simply launched found that companies run 53% in their workload in the cloud, and that 50% of the statistics that groups shop is likewise saved on the cloud [2]. It's far predicted that through the yr 2025, eighty percentages of businesses might

Doi:10.48047/ijiee.2025.15.5.32

have moved their operations to the cloud [3]. The overall amount spent on public cloud services is predicted to attain $304.Nine billion in 2021, that is a boom of 18.Four% over the amount spent in 2020 ($257.5 billion). This estimate becomes made by way of Gartner [4]. Computing in the cloud has stimulated the creation of a outstanding variety of apps which are geared at simplifying and enhancing many elements of human life. Some of those makes use of consist of online studying, on-line shopping, on-line entertainment, and online hospital treatment [5-15]. It gives 3 one-of-a-kind varieties of offerings: software as a provider (SaaS), platform as a carrier (PaaS), and infrastructure as a carrier (IaaS), in order that it is able to meet the necessities of a huge variety of users (IaaS). This era is extra famous than conventional infrastructure-based structures [16-18] because to the blessings inclusive of on-call for carrier, availability, scalability, and pay-as-you-pass pricing. Regardless of this, a number of troubles and limits need to be addressed, some of which consist of safety and privateness, aid control, elasticity, reaction time, and many others [19, 25].

The maximum essential trouble that should be solved for this generation to be broadly used is protection. Further to standard protection desires, the dynamic and scalable nature of the cloud provides positive specific security troubles. Furthermore, customers are not aware about the precise location of their records considering the fact that it's miles managed and kept remotely. Users need to therefore totally depend upon the cloud company to shield their information's protection and privateness. Therefore, it turns into vital for cloud vendors to keep security and privateness, considering that doing so might save you the enormous use of cloud computing. Identity and get entry to management, virtual signature, encryption and key management, assault detection and prevention, digital environment security, interface protection control, records garage safety, hardware protection, guarantee and compliance measures, and assault detection and prevention are all approaches that the cloud is made comfy (as proven in discern 1). This have a look at, however, makes a speciality of leveraging danger detection and prevention to secure cloud computing. DoS/DDoS attacks are substantial security breaches which might be used to intervene with cloud service availability. DoS assaults make about 22% of cloud assaults, claims the studies [50]. DDoS attack times are becoming extra common each year. Four.83 million DDoS assaults are recorded for the duration of the first half of of 2020, which is 15% extra than in 2019 [53]. Through 2023, there can be 15.4 million assaults worldwide, predicts Cisco [54].

Therefore, it's vital to create powerful defences in opposition to DDoS attacks. But, growing a DDoS attack detection system for the cloud is a tough and complicated venture. There are

Doi:10.48047/ijiee.2025.15.5.32

some of strategies for protecting in opposition to these assaults in a cloud surroundings, however none of them can accurately become aware of those assaults, leaving room for the improvement of new DDoS attack detection strategies. So that you can assure the provision of cloud offerings, this thesis investigates the many elements of DDoS assault detection for cloud computing. This study has a look at tries to solve the hassle of cloud computing safety. On this sense, the attempt specializes in growing defences in opposition to DDoS attacks. It also carries studies completed to signify DDoS assault detection strategies for cloud systems. This thesis provides gadget learning-based totally defences against DDoS assaults. Best few varieties of DDoS attacks are covered on this thesis. Moreover, this thesis solely uses simulation-based totally studies to assess the security answers that are cautioned. This thesis most effective offers a few solutions to the numerous problems in cloud safety. The primary contribution of this text consists of the design and development of machine getting to know enthused IDS to mitigate the impact of DDoS assaults in cloud environment.

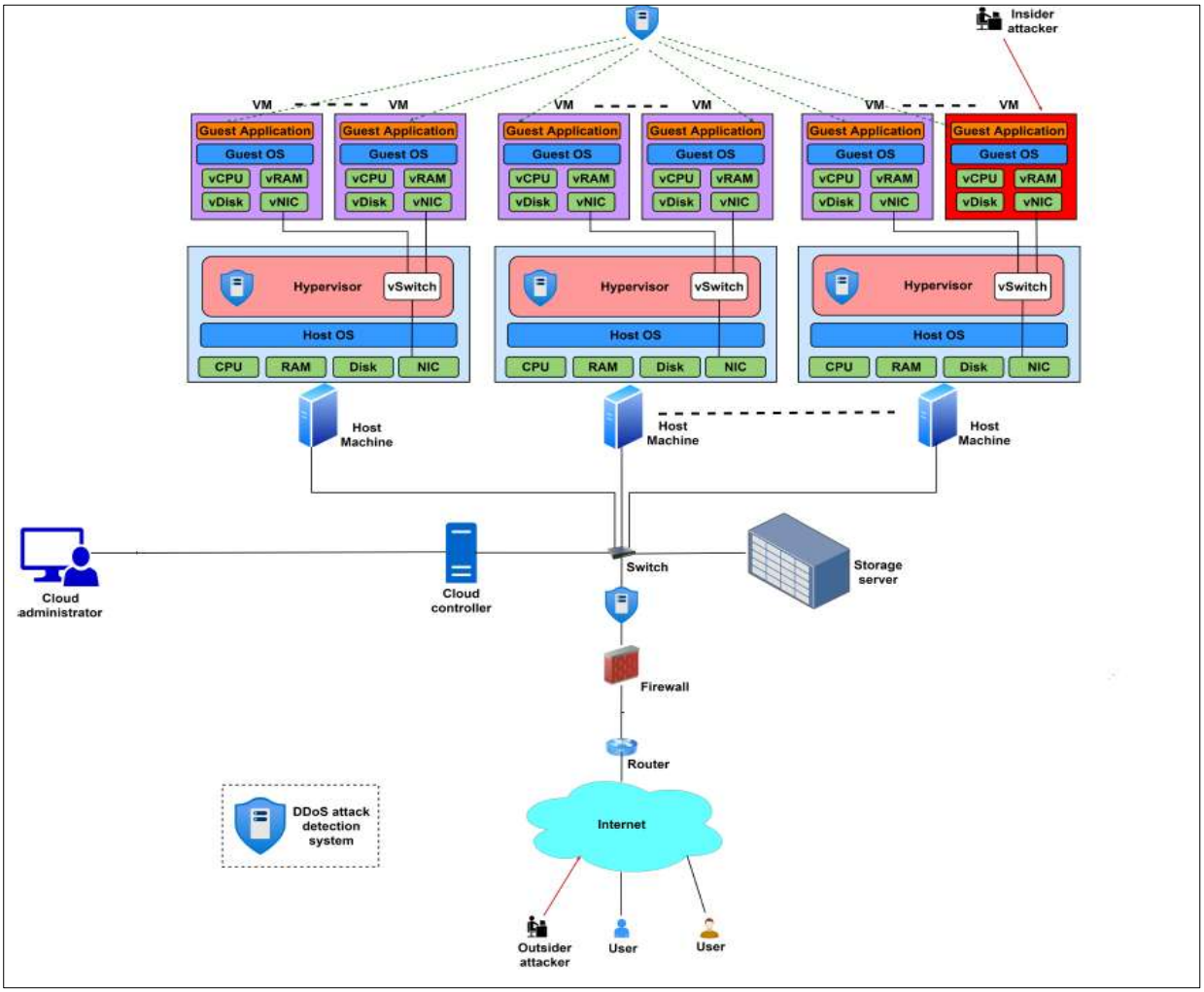


Figure 1: A conventional Cloud safety structure for DDoS assaults

Doi:10.48047/ijiee.2025.15.5.32

2. Related work:

Signature-based intrusion detection systems perceive on-going suspicious hobby by matching them to known dangerous patterns or commands. It keeps a file of the numerous different assaults and dangerous behaviours in a signature database that it continues. This signature database needs to be up to date on a normal foundation to be able to pick out lately released assaults. As a way to identify potentially dangerous behaviour on a community, current network packets are compared with a formerly stored set of regulations. Laugh, which was brought via Martin Roesch, is each one of the best and most drastically used signature-primarily based approaches (2015). Chortle is a extensively popular signature-based intrusion detection machine (IDS) which could capture packets and reveal network visitors in actual time. Parent three illustrates its primary additives, which include a packet decoder, a pre-processor, a detection engine, a logging and warning machine. The currently flowing traffic is pre-processed first, and then it is despatched directly to the detection engine. When pre-processing records, reproduction and incomplete data is removed. The present day packet is then compared with data that are stored in the signature database by means of the detection engine. If there may be a healthy, then the alert is produced for the applicable authority; otherwise, the packet is despatched as a regular packet. If there is no healthy, then the alarm isn't always generated (Roesch, 2015).

A big range of research furnished signature-based IDS, inclusive of the only that Lin et al. (2012) evolved, so that you can perceive known threats in a cloud context. Records are accumulated from the running systems of all the VMs and then dynamically up to date in order that the detection standards can be configured. Of their 2010 examine, Lo et al. delivered a cooperative intrusion detection structure (CGA). Each server has an intrusion detection gadget (IDS) that may be a mix of a signature database and a block desk that maintains a document of preceding assaults. The packet is first compared to the block desk, and then it's miles as compared to known signatures. Current attacks have a more opportunity of being gift, thus it's miles critical to take a look at them first. A threshold value is used to decide the severity of an peculiar packet thru transfer alert clustering, which presents statistics approximately the aberrant packet. The intruder detection device (IDS) then discards the malicious packet. In a similar vein, Meng et al. (2014) suggested strategies for signature-based IDS. In a situation together with opposed network site visitors, the authors state that the hazard of a mismatch is

Doi:10.48047/ijiee.2025.15.5.32

plenty better than the probability of a suit taking place. As a result, they identified the assault by means of using a coverage known as mismatch. Of their 2015 observe, Mandal and associates supplied a signature-based totally IDS technique to pick out application-stage threats. This technique involves putting a sniffer among the consumer and the cloud provider. The sniffer then collects the data packets and sends them to the parser in order that they'll be similarly processed. The output of the parser is as compared to the accumulated semantic rules by way of a parsing grammar, and the result is fashioned based totally on this comparison.

Notwithstanding the truth that a signature-based totally IDS can pick out recognised threats quickly and has a totally low prevalence of fake positives, it wishes ordinary updating of its signature database. Anomaly detection techniques were created a good way to ward off the limitations imposed by using signature-based intrusion detection systems. This method creates a behavioural profile through studying the user's conduct in an effort to accomplish that. After then, this behavioural profile is applied to stumble on attacks, whether they're recognized or unknown. A essential example of the working idea for an anomaly detection technique is shown in determine four. It is created from primary elements, particularly the phase of education, and the section of detection. During the education phase, the function introduction module gathers input from the host device or community and pre-processes it in order that capabilities can be built. The training module will utilize those trends to assemble a behavioural version of the situation. The facts are classified according to this version as either every day or aberrant (intrusion) behaviour. This version is what is used to identify intrusion all through the ambiguity detection section. Any variant from the everyday float of visitors is seemed as an intrusion, and an alarm is precipitated inside the gadget that indicators the security administrator (Sari, 2015). This approach can perceive fresh threats, however it makes heavier demands on the computer's processing assets. Due to the fact each departure from common conduct triggers an alert, it's miles now the responsibility of the safety supervisor to decide the motive of the alarm. The category of anomaly-primarily based procedures is in addition damaged down according to the technique that is used to discover anomalies. A few examples of such methods are gadget gaining knowledge of, fuzzy good judgment, support vector machines, and records mining. In the latest decade, some of studies have investigated using those strategies. To create their behavioural version, Kumar et al. (2011), as an example, made use of a hidden Markov approach. This method uses a log record to monitor the frequency of gadget calls so that you can become aware of probably dangerous activity. An IDS is constructed with the use of 3 different profiles, known as low, medium, and excessive,

Doi:10.48047/ijiee.2025.15.5.32

each of which correlates to distinct elements of current hobby. The low profile is used to depict styles that have a bit hazard of being matched correctly. The excessive profile, then again, refers to styles that have a totally high probability of being matched. Ultimately, the profile that satisfactory fits the present is the one that falls within the center. Every profile is matched the usage of a threshold cost that has been described in advance.

Yuxin (2011) suggested a method that makes use of the system-mastering technology and is based on the static application behaviour evaluation. It operates in ranges: the primary level is to decode programs, and the second one level is to build context-free grammar in order to mirror the waft of the manner. We check out and piece collectively all the branches in order that we can also acquire the complete sequences. The whole of the system calls had been condensed into some concise sequences. They used awesome methods for deciding on functions: one called statistics benefit (IG), and the alternative was called record Frequency (DF). An IDS technique the usage of a -tier machine became evolved by means of Srinivasan et al. (2012). This method is a hybrid of unsupervised mastering and supervised type. For the reason of intrusion detection, Wolthusen (2012) used the frequency of regular/ordinary machine calls. To begin, all through the course of sometime length, a huge quantity of information is gathered from each VM. The methods make the idea that the VMs are not going to be malicious for a certain amount of time as soon as the start-up process is whole. It has a temporal complexity of O (n), wherein n is the full range of lines inside the expression. This technique has a detection fee of 100%, with simply 11 percentages being fake positives.

3. Proposed method:

IDSs are taken into consideration as a becoming device for checking and spotting facts move, and that they give security in opposition to interruptions that compromise the accessibility, trustworthiness, and privateness of a records framework. The company interruption identification system for allotted computing weather that has been introduced in this text is completed in light of the extraction of highlights and resolution of these factors making use of a democratic based totally outfit classifier. The CIC-DDOS 2019 dataset fills in because the contribution for the examination and has been pre-handled utilizing encoding, scaling, and cleansing procedures. From that point ahead, the dataset that has been standardized is adjusted utilising the Destroyed methodology, which includes oversampling. Making use of RNN (intermittent brain community roused) t-Circulated Stochastic Neighbour implanting issue, the factors are recovered from the decent dataset. The most worthwhile highlights are selected with

Doi:10.48047/ijiee.2025.15.5.32

the aid of math streamlining utilized hereditary calculation (GAAOLGA). In addition, the picked highlight units are then partitioned into a practise dataset and a testing dataset. These attributes are arranged utilising the troupe casting a poll classifier approach that became encouraged. Woodpecker and Flamingo Inquiry advancement (WPFS) coordinated meta heuristic calculation is brought in this article to work at the exactness of the characterization.

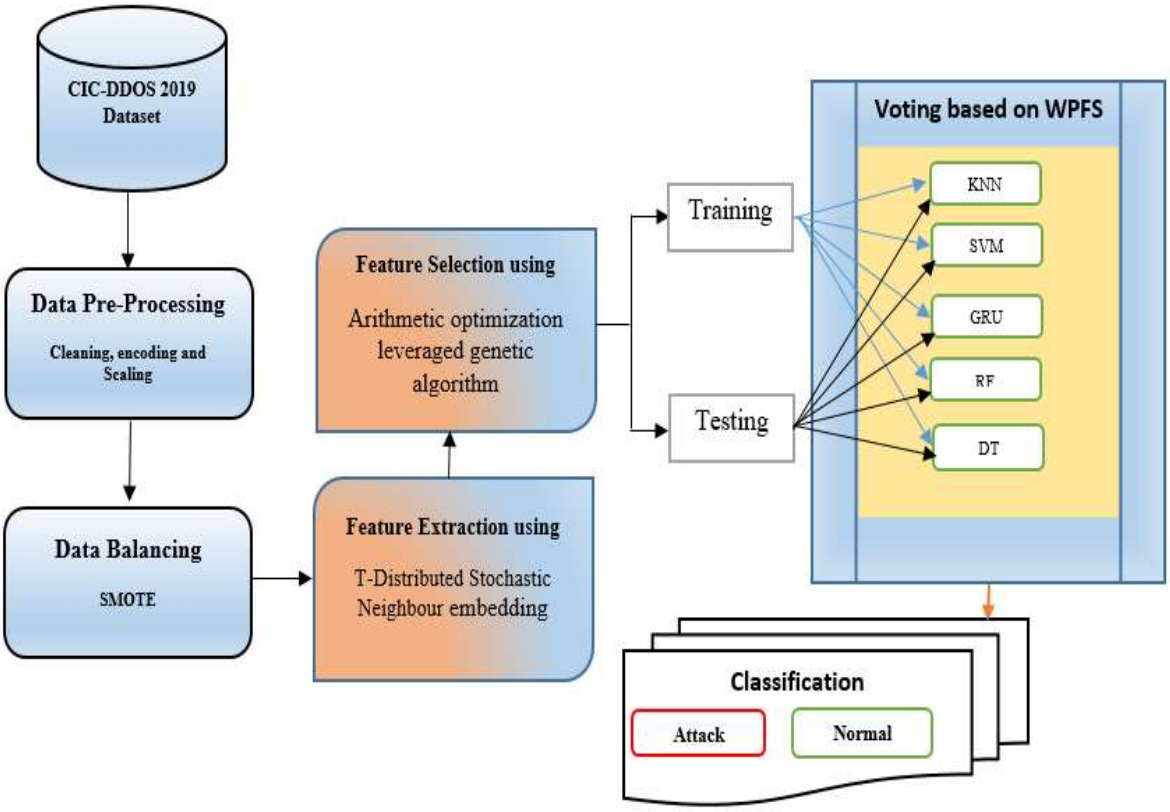


Figure 2: Schematic Representation of the Proposed Hybrid Framework.

3.1 Data prepossessing:

Because the dataset includes category values, numeric values, redundant and copy information, the proposed answer utilizes the pre-processing task. Approaches like as statistics cleansing, facts encoding, scaling, and normalization are carried out at the raw records earlier than the best analysis can be crafted from them. The technique of information cleaning includes removing any facts that is redundant, irrelevant, incomplete, or erroneous before it's miles used for evaluation. This facilitates the facts turn out to be ready for analysis. The wrong layout produces findings which can be faulty whilst the use of the standard facts. The records cleansing topology improves accuracy without disposing of any facts this is required. Due to the fact the laptop can handiest study numerical data, the categorical values of the CIC-DDOS

Doi:10.48047/ijiee.2025.15.5.32

2019 dataset are translated into numerical values. Consequently, all symbolic representations of traits are modified over to their corresponding numerical values for the duration of the procedure of records encoding. However, the overall performance of the classifier is negatively impacted due to the multiple scaling's used for the capabilities. It's miles essential to normalize the information for the reason that traits may have very huge numerical values. The traits are fixed using normalization to have a numerical fee between 0 and 1, which would possibly variety from zero to one. With a purpose to normalize the records, the Min-Max method, which is the most effective and maximum truthful choice, is utilized in Eqn (1). A unified data format is to be had for the dataset after cleaning, encoding, and normalization have been finished.

$$\mathrm{K}:\mathrm{K}_{nor}=\frac{K-K_{\min}}{K_{\max}-K_{\min}} \tag{1}$$

The dataset feature space (K) is signified by $U(\mathrm{K}_1,\mathrm{K}_2,....,\mathrm{K}_n),1<i<I$. In which the entire wide variety of instances is denoted as *I* within the feature space.

3.2 Data Balancing:

Because the class imbalance issue may additionally have an effect at the pre-processed functions, the SMOTE scheme is used if you want to discover a approach to the hassle. Discovering an instance's okay-nearest neighbours is one of the obligations that the SMOTE technique in [33]. After that, with a view to generate a line section inside the function area, it chooses at random an additional k-nearest neighbour to link with the primary neighbours. Within the method of finding the ok-nearest neighbours, the Euclidean distance is used to kind the diverse instances. Ultimately, each example is reset to the SMOTE class that allows you to produce artificial times. Synthetic instances are created commonly for the reason of computing the difference among the original instance and its closest neighbour example, multiplying that distinction with the aid of a random integer starting from 0 to at least one, after which including the result to the unique instance. As a result, the problem of instances being over fit has been averted way to this. The considered dataset has a as an alternative massive wide variety of ordinary occurrences whilst as compared to the variety of attack times. If the dataset incorporates training which might be imbalanced, the accuracy of the type will go through as a end result. As a end result, the SMOTE approach is used to the dataset as a way to bring it into higher stability. Artificial examples are produced if you want to be introduced to the

Doi:10.48047/ijiee.2025.15.5.32

preliminary dataset, and then the classification version is skilled the use of samples from the brand new dataset.

3.3 feature extraction using t-distributed stochastic neighbour embedding mechanism (t-SNE):

This segment introduces t-SNE set of rules for characteristic extraction based on hierarchical deep neural community. There are ranges concerned within the t-SNE set of rules. Inside the first section of the technique, a probability distribution is performed in a area with a high measurement. As a result, the benefit with which a pair of items within the space may be picked is directly proportional to the degree to which they resemble each other. Alternatively, the probability of picking extraordinary styles of matters is reduced. Similarly, the possibility in a low-dimensional area is generated, and the excessive-dimensional opportunity distribution is corresponding to the low-dimensional probability distribution. Whilst calculating similarity distances, many methods each appoint their personal precise set of criteria (which include okay-approach using Euclidean distance). In an effort to illustrate the similarity distances among two items, the t-SNE technique relies on conditional chance.

In this context, the pre-processed dataset ‘y’ which is given as an input for with N samples $y = (y_{1,} y_{2,} y_{3,} ..... y_{n,})$ the distance between any two samples y_i and y_j will be determined based on Eqn.(2) & (3).

$$Q_{ij} = \frac{Q_{i|j} + Q_{j|i}}{2N} \tag{2}$$

$$Q_{j|i} = \frac{\exp(-\frac{|| y_i - y_j ||^2}{2\sigma^2_i})}{\sum_{k \neq 1} \exp(-\frac{|| y_i - y_k ||^2}{2\sigma^2_i})} \tag{3}$$

3.4 Feature Selection Using Genetic Algorithm Leveraged Arithmetic Optimization (GAAO)

On the subject of defensive the sensitive data in cloud computing networks, low-complexity and excessive-accuracy intrusion detection measures are vital. As a end result, the machine getting to know-primarily based approach improves intrusion detection overall performance, and the detection rate is raised via the feature selection-primarily based, nature-inspired

Doi:10.48047/ijiee.2025.15.5.32

optimization approach. The CIC-DDOS 2019 dataset has a huge range of characteristics, a number of which might be useless, and a massive range of capabilities that have been extracted from the dataset. The effects show that the IDS classification overall performance is better with the aid of the optimized function choice approach. That is wherein the GAAO set of rules is available in to select the first-rate characteristics. Despite the fact that the advised function selection approach ranks capabilities in keeping with their usefulness, it could not become aware of the simplest capabilities for training a classifier. The advised method begins by using sorting all traits into businesses according to their usefulness inside the category methods. The functions that had been retrieved are break up into a schooling set and a test set. The GAAO algorithm is used to examine the education records and determine which characteristics are maximum useful. The classification algorithm is fed the quality functions extracted from the schooling and test data so as to examine the version's success. A good way to reduce the error delivered through every generation at the same time as concurrently improving category accuracy, Equation (four) fashions the perfect function choice.

$$obj\_g(x) = \min E(y) \tag{4}$$

$x = |y|,\ y \subseteq Y$ Where, Denotes the gathering of authentic traits (Y) selected so that you can attain the excellent feasible objective feature.

The capabilities are created in a subset at each GAAO populace instance. For the health calculation, the selected functions are fed into classifiers like SVM kernel, KNN, Bi-LSTM, GRU, and RF. In such models, the most beneficial capabilities are selected by using a minimization of the fitness function. Equation (5) defines mistakes because the discrepancy between actual $h(k)$ and predicted output $\hat{h}(k)$.

$$N(k) = h(k) - \hat{h}(k), \qquad k = 1, 2, ...n$$

Where k stands for experimental records. Equation (5), frequently called the health characteristic, calculates the fraction of the whole errors by way of the full wide variety of observations.

$$fitness(i) = \frac{\sum_{i=1}^{n} N(k)}{n} \tag{5}$$

Doi:10.48047/ijiee.2025.15.5.32

First-class and worst fitness are used to decide the minimum and maximum fitness, respectively. By fixing Equation, the local quality price is determined for the duration of the exploratory seek the use of equation 6. In its pursuit of the fine solution, the GAAO set of rules continuously adjusts the parameters of Equation 7. In this step, the people who constitute the minimal health answer are picked to preserve into the following technology. This technique gradually improves the classifier through along with new traits. As soon as the first-class viable class accuracy has been attained inside the schooling dataset, the very last choice is made at the top-quality quantity of functions to use in the technique. Subsequently, a feature set is settled upon for the proposed GAAO function choice technique.

$$y_{i,j}(\mathrm{y}+1)=\begin{cases}\mathrm{a}(\mathrm{y}_j)\div(MOA+\in)\times\left(\left(u_j-l_j\right)\times\mu+l_j\right), & r_2<0.5\\ \mathrm{a}(\mathrm{y}_j)\times MOA\times\left(\left(u_j-l_j\right)\times\mu+l_j\right), & otherwise\end{cases} \tag{6}$$

Various sections of the quest space have been arbitrarily explored the usage of the department (D) and Multiplication (M) techniques to locate higher answers. The version may be said using the equation (5). Throughout this time of looking, the MOA characteristic is what determines the parameters. $r_1 < MOA$. The operator (D) is conditioned by $r_2 < 0.5$ and ignored by the operator (M) until the current task is completed. Where, $y_i(\mathrm{y}+1)$ is represents the next iteration i^{th} solution, $y_{i,j}(\mathrm{y}+1)$ represents the current iteration i^{th} solution in j^{th} position, and best obtained solution in j^{th} position is denoted as $b(x_j)$. Then the j^{th} position upper and lower bound value is denoted into u_j and l_j, respectively. The random numbers are denoted as r_1 and r_2, respectively. The search process is adjusted by the control parameter μ, the value is 0.5.

To perform the exploitation looking section, the operators S and A do a subtraction and an addition. Furthermore, with the extra stipulation that the present cost of the asset isn't always greater than the MOA(*y*) feature, the exploitation will only arise. S and A operators carry out in-intensity search on numerous excessive-density locations in GAAOA, in the end main to an appropriate solution described by using Equation (7).

$$y_{i,j}(\mathrm{y}+1)=\begin{cases}\mathrm{a}(\mathrm{y}_j)-(MOP\times\left(\left(u_j-l_j\right)\times\mu+l_j\right), & r_3<0.5\\ \mathrm{a}(\mathrm{y}_j)+MOP\times\left(\left(u_j-l_j\right)\times\mu+l_j\right), & otherwise\end{cases} \tag{7}$$

Doi:10.48047/ijiee.2025.15.5.32

Math Optimizer Probability (MOP) is a coefficient, and the value of the function at each new release is denoted as $MOP(y)$, and the sensitive parameter is denoted as α.

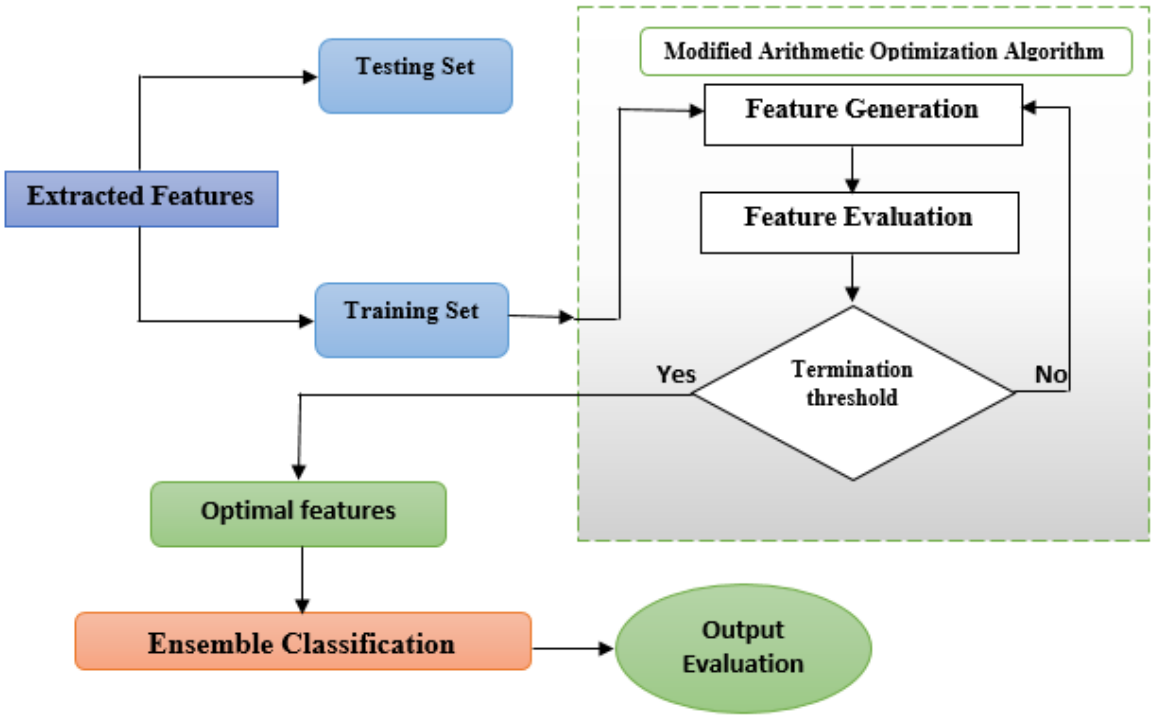


Figure 3: Proposed characteristic choice using Genetic set of rules leveraged mathematics

Pseudo code for proposed Genetic algorithm leveraged Arithmetic Optimization

Initialize: η, λ *and* $j = 1, 2...., N$

While $(y < \max(y))$ ***do***

Evaluate fitness and select best solution.

Compute $MOA(y) =, \min + y \times \left(\frac{\max - \min}{\max\ y} \right)$ *(8)*

Compute $MOP(y) = 1 - \frac{y^{1/\alpha}}{\max\ y^{1/\alpha}}$ *(9)*

for *(i=1 to features)* ***do***

Doi:10.48047/ijiee.2025.15.5.32

```
        for (j=1 to locations) do
            Generate random values between (0,1)
            if r2 > MOA then
                Discovery stage
                if r2 > 0.75 then
                    update the positions in Eq. (6) based on the dividend operator D
                else
                    update the positions in Eq. (6) based on the multiplication operator M
                end if
                Exploitation stage
                if r3 > 0.5 then
                    update the positions in Eq. (6) based on the subtraction operator S
                else
                    update the positions in Eq. (6) based on the addition operator A
                end if
            end if
        end for
    end for
    y = y + 1
end while
Return best
```

3.5 Ensemble Voting Classifier based on Woodpecker based Flamingo Search optimization:

For the reason of community intrusion detection, that is used to study the viability of the techniques, ensemble studying gives a number of characteristics that make it superior than a unmarried classifier. Following the function selection method, the training facts are then sent to the classifier to learn. A vote casting classifier-based totally hybrid optimization approach is used inside the recommended method so that you can categorize the community traffic that is being monitored. For you to improve the accuracy of the ensemble classifier, the wooden Peaker Optimization algorithm (WPOA) is merged with the Flamingo search Optimization (FSO). These ensemble gadget mastering classifiers use a aggregate of SVM kernel, KNN, Bi-

Doi:10.48047/ijiee.2025.15.5.32

LSTM, GRU, and RF. Those 4 wonderful machine getting to know classifiers, each of which is primarily based on arithmetic optimization, are skilled the usage of the specified functions. The balloting method of the proposed hybrid WPOA with FSO set of rules, that's known as WPFSO, optimizes the weight characteristic of each classifier so as to gain greatest performance (wooden Peaker based totally flamingo seek optimization). The cause of the ensemble classifier is to beautify category accuracy even as concurrently lowering error fees. As a consequence of the fact that individual classifiers are able to generating varying results, the ensemble classifier draws from the advantages provided via man or woman classifiers to achieve more accurate predictions.

On this hybrid technique, an exploration technique is used to improve the general best answer with each cycle of WPOA evaluation. The goal of Hybrid WPFSO is to make the answer higher by retaining the worldwide excellent role up to date. If there is no development within the number of WPOA cycles, the previous first-class answer might be switched out for a new one that became calculated using the FSO technique. This new perfect stage in reality transmits a sign to the mathematics operator, which suggests the math operator who turned into the previous first-rate solution might be brought on to adjust its path in reaction to this signal. The blessings of each the WPOA's capability to discover and the FSO's belongings in terms of nearby search are blanketed into the WPFSO set of rules. Although WPOA does now not exhibit improvement at the high-quality factor, FSO provides to the method of upgrading the first-rate answer globally. Similarly to that, this may take region several instances at some point of the manner of finding a solution. The algorithm is kicked off with WPOA via developing a population and primary setting the parameters for each strategies. In WPOA, the FSO set of rules is used for the computation of the health solution in the course of each iteration. If there may be no development made to the current great answer all through a cycle, then the FSO will use this best solution as the starting point for computing the fine place. It will be carried on until the point whilst termination takes area. After then, the procedure for searching is done yet again with WPOA using the now finest solution similarly to the previously finest solution. This manner is repeated till the convergence requirements had been happy. The balloting based totally WPFSO algorithm that has been suggested can be done by means of following those steps.

Algorithm: Voting based WPFSO algorithm
Input : Various ML (Machine Learning) MODELS , DDOS Optimal Attack Features

Doi:10.48047/ijiee.2025.15.5.32

Output: Selection of best ML Model based on the Optimal Attack Features

Step 1. *Initialize* $\begin{bmatrix} V_{data} \leftarrow 0; \\ v_{index} \leftarrow 0; \\ T_a \leftarrow 0; \end{bmatrix}$ *# voting data, voting index and total accuracy are initialized to '0';*

Step 2. *Initialize population of WPFSO based on the attack features*

Step 3. *Select randomized inputs from the attack features let us say if No. of feature F>1 then compute fitness function.*

Step 4. *If identified features < 3*

Step 5. *model←model.fit(train with all the parameters)*

Step 6. *Endif*

Step 7. *for model in MODELS do*

Step 8. *model←model.fit(train with all the parameters)*

Step 9. *model←model.predict (X_Test)*

Step 10. *accuracy ⇐ metrics.accuracy_score(Y _Test, predict)*

Step 11. *VotingData ⇐ VotingData + predict ∗ accuracy*

Step 12. *TotAccuracy ⇐ TotAccuracy + accuracy*

Step 13. **end for**

Step 14. *Check for improvement in the global optimal solution. If (* P *in iteration* $i+1$ *is not less than* P *in iteration* i *,*

Step 15. *If the random number is not less than* p *then move randomly* $F_i^{t+1} = \left(r^2 \times a_j^t - a_k^t\right) \times pF_i$

The present best solution is denoted as (b^*)*, solution vector is denoted as* F_i^{t+1}*, iteration number, random number and fragrance is represented as* t, r *and* F_i*, respectively.* a_i^t *is used to update the butterfly position during iterations.*

Step 16. *The global optimal solution is obtained then goes to step 3 and repeat steps until the termination criteria will met.*

In the WPFSO technique that has been provided, the first-rate feasible outcome is continually given to the output after each execution and diagnosis of a new instance. In end, the output that

Doi:10.48047/ijiee.2025.15.5.32

is ideal is finished at the same time as maintaining the greatest possible accuracy and experiencing the fewest possible errors.

4. Results and Discussions:

Which will validate the proposed technique, the CIC-DDOS 2019 dataset become used. Experiments are finished on a laptop with a 2.3 GHz Intel center i3 CPU, and the MATLAB R2020a tool is used for both implementation and evaluation of the experiments. The synthetic dataset utilized in CIC-DDOS 2019 [13] become generated with the assist of the freeware utility NetSim. This device is capable of CC, LTE networks, Cognitive radio networks, and a spread of other community sorts. The gateway, stressed nodes, sensor nodes, and routers are the additives that make up the CC community. Following the engagement of the captured packets in their respective CSV distinct documents for every assault, the aforementioned files are then merged into the CIC-DDOS 2019 dataset. It has twenty one-of-a-kind numbers of functions and two label features further to those. Table 1 includes a listing of the dataset's traits, even as the dataset itself contains seven distinctive kinds of attacks, such as Sybil, Clone identification, hello flooding, local restore, Selective forwarding, Sinkhole, and Blackhole.

Table 1: Feature set utilized from the dataset

Id number	Feature name
3	Control Packet Type
4	Source ID
5	Destination ID
6	Transmitter ID
7	Receiver ID
8	APP Layer Payload
9	TRX Layer Payload
10	NW Layer Payload
11	MAC Layer Payload
12	PHY Layer Payload
13	PHY Layer Overhead
14	APP Layer Arrival Time
15	TRX Layer Arrival Time
16	NW Layer Arrival Time

Doi:10.48047/ijiee.2025.15.5.32

17	MAC Layer Arrival Time
18	PHY Layer Arrival Time
19	PHY Layer Start Time
20	PHY Layer End Time

The dataset is partitioned into categories according with the breakdown shown in table 2. As is standard, the dataset is partitioned, and successful attacks are counted in the direction of trying out and schooling persistence. There have been 42,587 total incidents of attacks.

Table 2: Category wise dataset description

Category	Counts
Clone_ID	5852
Hello_fooding	5854
Local_repair	5858
Selective_forwarding	5625
Sinkhole	5898
Blackhole	5454
Sybil	5255
Normal	53696

Table 3: Testing and Training dataset description

Label	Training dataset	Testing dataset
Attack	42227	17581
Normal	125689	54656

4.1 performance Metrics:

The performance of the proposed NIDS became evaluated the usage of among the matrices which might be discussed further down. The goal of this have a look at based totally on NIDS is to attain the best feasible stage of accuracy in predicting occurrences from the test dataset. Accuracy, real high-quality (TP), fake nice (FP), genuine bad (TN), false negative (FN), bear in

Doi:10.48047/ijiee.2025.15.5.32

mind, Precision, and F1 rating are the metrics which are used to assess the overall performance. Equation 10 [36] is what's used to decide how accurate something is.

$$A = \frac{TN + TP}{FP + FN + TP + TN} \qquad (10)$$

In this scenario, TP is described as the proportion of assaults which might be properly diagnosed, FP as the proportion of attacks that are mistakenly diagnosed, TN as the percentage of attacks which are efficiently identified, and FN as the proportion of assaults which can be inaccurately diagnosed.

$$precision(p) = \frac{TP}{TP + FP} \qquad (11)$$

$$recall(r) = \frac{TP}{TP + FN} \qquad (12)$$

$$Fscore = 2\frac{p \times r}{p + r} \qquad (13)$$

4.2 Balancing Data Using SMOTE:

There are fewer assault classes to be had than there are trendy elegance slots to be had. As a result, the dataset has to be balanced, therefore the oversampling technique of SMOTE is applied. In the SMOTE technique, the nearest neighbor parameter value okay=four is used to use an oversampling method to the facts in order to normalize it and make it comparable to different lessons. Due to the performance of SMOTE, the variety of attack instances has been raised to 142586 capabilities. The performance of the SMOTE contributes to an development within the accuracy of the NIDS system. Inside the absence of SMOTE stability, the precision is subpar, and the margin of error is good sized. Desk four provides the results of a evaluation among the overall performance with and without balancing. Primarily based on table four, the advised technique has obtained an accuracy of 98% while balancing, while the accuracy drops to ninety five% whilst balancing is not carried out.

Table 4: Comparison of SMOTE balancing and without balancing performance

Parameters	**With balancing**	**Without balancing**

Doi:10.48047/ijiee.2025.15.5.32

Accuracy	0.9849	0.9564
Error	0.0215	0.0482
Sensitivity	0.9709	0.9283
Specificity	0.9867	0.996
Precision	0.9747	0.9284
FPR	0.1057	0.0986
F1_score	0.9723	0.9271
MCC	0.9701	0.9216
Kappa	0.9193	0.7926

Features are extracted from the dataset using the t-distributed Stochastic Neighbour embedding mechanism approach after the dataset has been balanced. There are simply 14 traits to work with before extraction, but thereafter there are 365. The function choice approach chooses the maximum useful traits from among the ones that have been retrieved.

4.3 Feature Selection Using Genetic algorithms leveraged Arithmetic Optimization:

The GAAO algorithm is used for you to pick the capabilities within the handiest way possible. This permits for improved function selection. The health feature is used to pick the first-class features, which means that a low error fee shows the first-class characteristic set from the dataset. Parent three presentations the convergence curves of the counselled technique, which demonstrates that AOA has a superior capacity for exploration. The GAAO technique has reached the factor wherein the minimal average blunders fee of zero.0183 has been mounted on the way to choose the ideal functions. Whilst contrasted with GA, PSO, SHO, and CSA, in addition to ALO, the feature choice algorithms with the least amount of errors for the present scenario are indexed in descending order from pleasant to worst. The accuracy of function choice has been progressed due to the fact to the GAAO approach that changed into provided. The set of rules's parameters are detailed in table 5.

Table 5: GAAO Parameters

GAAO	
Parameters	**Values**
Search solutions	36

Doi:10.48047/ijiee.2025.15.5.32

Maximum iteration	125
MOP max	1
MOP	0.23
μ	0.633
Lower and upper bound	0 and 1

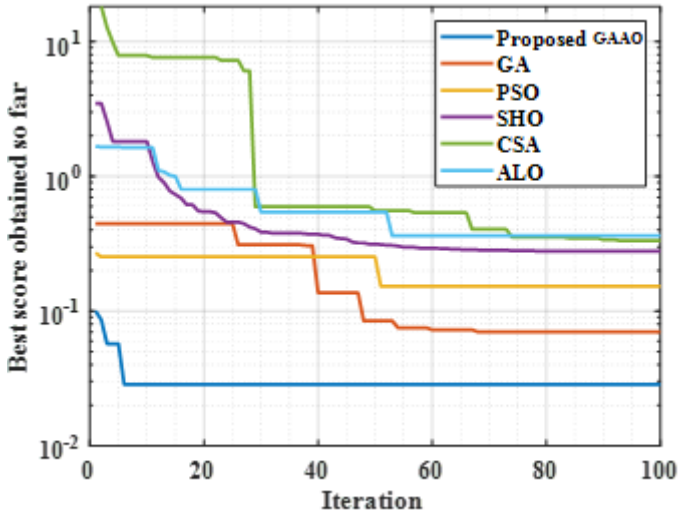


Figure 4: Convergence Comparison

4.4 Performance Analysis of Voting Based Classifier:

An ensemble classifier that makes use of the SVM kernel, kNN, RF, Bi-LSTM, and GRU was recommended for use with IDS. The couple of classifier version has blanketed a voting machine primarily based on majority rule on the way to choose the most useful direction of action. While it got here time to make the very last forecast, a weighted majority vote changed into used to mixture all of the preceding guesses approximately how the class would turn out. A weighted version of the records was used to broaden a education sequence that became then added in a sequential style to primary rookies. As the basis classifier, the k-nearest neighbours' set of rules is used. The WPFSO algorithm's hybrid method turned into used to best the optimization of the burden characteristic. Times of ordinary attacks in the dataset are given the weights that have been optimized for them. Normal assaults within the dataset are then given the most efficient weights. Performance of the proposed WPFSO set of rules is proven in parent 4. Comparisons to other algorithms such as GA, PSO, SHO, CSA, and ALO have

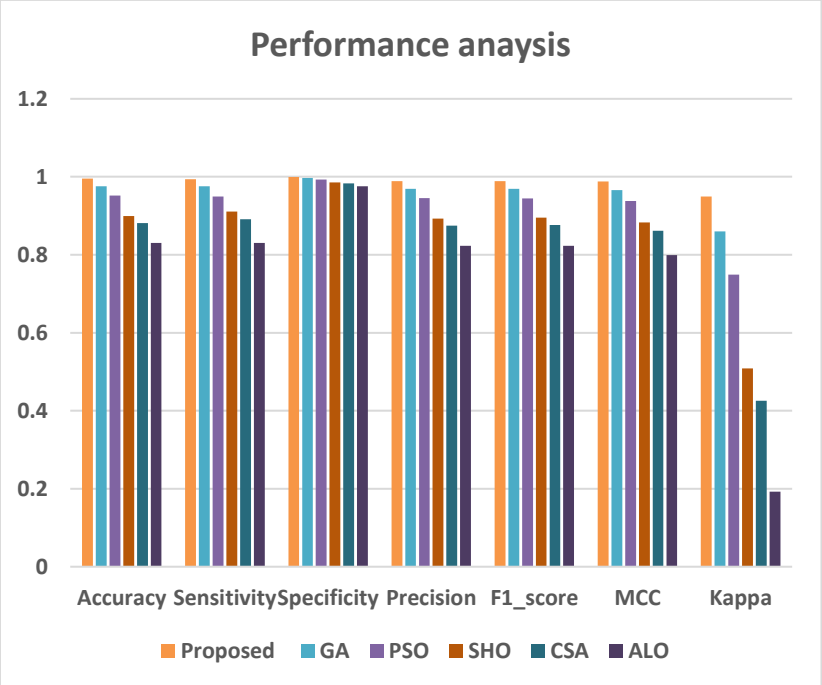
Doi:10.48047/ijiee.2025.15.5.32

proven the WPFSO algorithm's efficacy. The mean squared error (MSE) is a metric used to degree how properly a classifier performs with the aid of comparing the classifier's actual output to the preferred output. You may write it down as an equation (14),

$$MSE = \sum_{i=1}^{n} (o_i^t d_i^t)^2 \quad (14)$$

Table 6: Performance analysis of voting based proposed WPFSO algorithm

Parameters	Proposed	GA	PSO	SHO	CSA	ALO
Accuracy	0.9957	0.9761	0.9518	0.8992	0.8811	0.83
Sensitivity	0.9937	0.9754	0.9497	0.9109	0.8911	0.8305
Specificity	0.9996	0.9968	0.9933	0.9858	0.9832	0.9759
Precision	0.989	0.9694	0.9453	0.8926	0.8744	0.8232
F1_score	0.9889	0.9694	0.9447	0.8951	0.8767	0.8231
MCC	0.9881	0.9662	0.9379	0.8833	0.8619	0.7995
Kappa	0.9498	0.8601	0.749	0.5085	0.4259	0.1922
FPR	0.0016	0.0044	0.0079	0.0154	0.018	0.0253
Error	0.0115	0.0311	0.0554	0.108	0.1261	0.1772



Doi:10.48047/ijiee.2025.15.5.32

Figure 5. Performance analysis

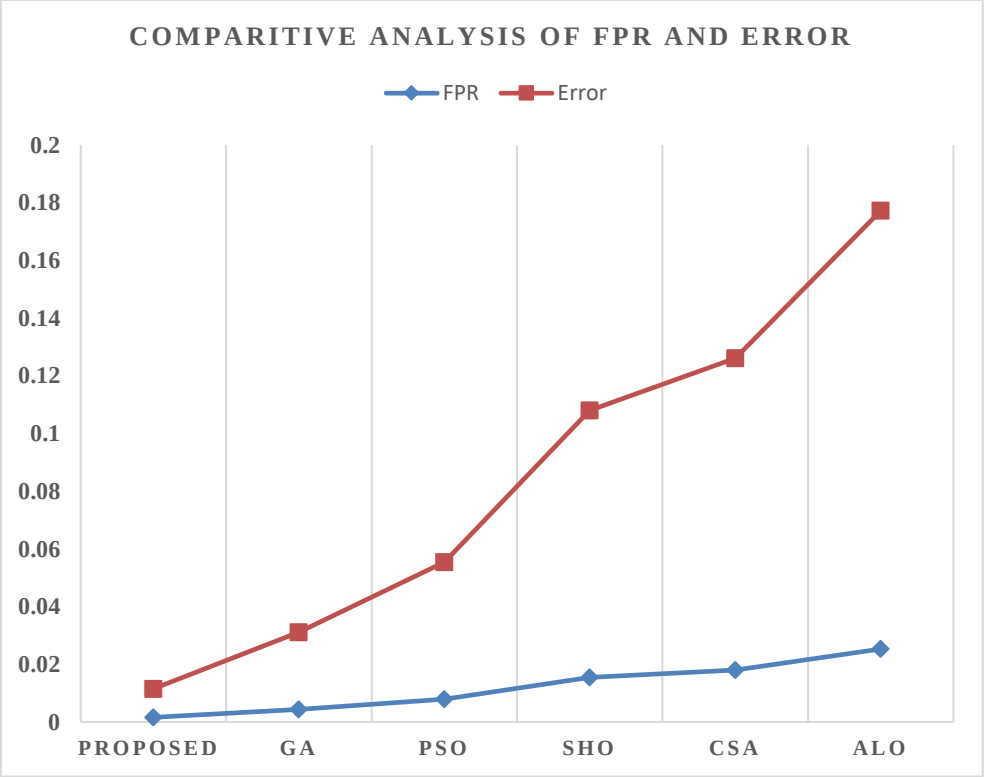


Figure 7: FPR VS Error Rate

The effectiveness of the proposed classifier is measured by way of its accuracy relative to those acquired the use of the person techniques (RF, kNN, SVM kernel, GRU, and Bi-LSTM primarily based ensemble classifier). The A-BO weighted majority balloting approach outperformed other methods at the dataset. A confusion matrix depicts the voting classifier's precision. Present vote casting algorithms together with GA, PSO, SHO, CSA, and ALO are in comparison to the proposed method, demonstrating its advanced accuracy. GA, PSO, SHO, CSA, and ALO balloting accuracy is 96.Seventy nine, 94.50, 89.Forty five, 88.Fifty three, and eighty two. Fifty seven percentage, respectively. The advised vote casting set of rules, but, has attained an accuracy of ninety eight.17 percentage, setting it ahead of the alternative techniques. The confusion matrix for both new and present day voting classifiers is proven in determine 6.

Doi:10.48047/ijiee.2025.15.5.32

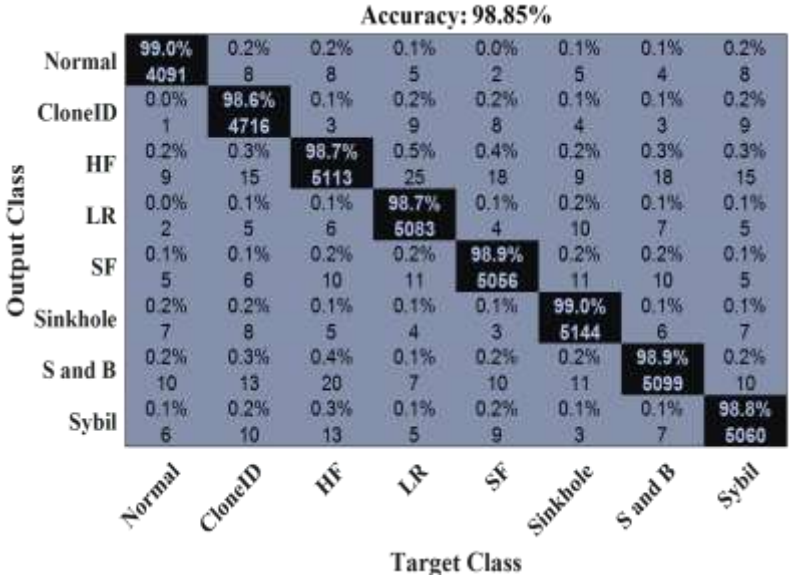


Figure 8. Confusion Matrix depicting the Performance of the proposed WPFSO Algorithm

The effectiveness of the proposed WPFSO balloting classifier scheme is likewise compared with other ensemble strategies. Adaptive boosting and bagging techniques are as compared with WPFSO balloting classifier to show the performance of the proposed set of rules. The mistake of the proposed technique (0.0321) is lower than the existing strategies Ada improve (zero.912) and bagging (0.078). Also, the performance of sensitivity, specificity, precision, FPR, F1_score, MCC, and kappa is outperformed the proposed technique in comparison to existing methods. Table 7 presentations the comparison of the proposed and present balloting classifier.

Table 7 Comparison of the proposed and existing voting classifier

Parameters	Voting WPFSO	Ada boost	Bagging
Accuracy	0.984	0.9473	0.9243
Sensitivity	0.953	0.9176	0.891
Specificity	0.9967	0.9936	0.9905
Precision	0.9596	0.9156	0.8835
F1_score	0.9564	0.9159	0.8856
MCC	0.9483	0.9052	0.8722
Kappa	0.8768	0.772	0.6671

Doi:10.48047/ijiee.2025.15.5.32

Error	0.0321	0.055	0.078
FPR	0.0045	0.0076	0.0107

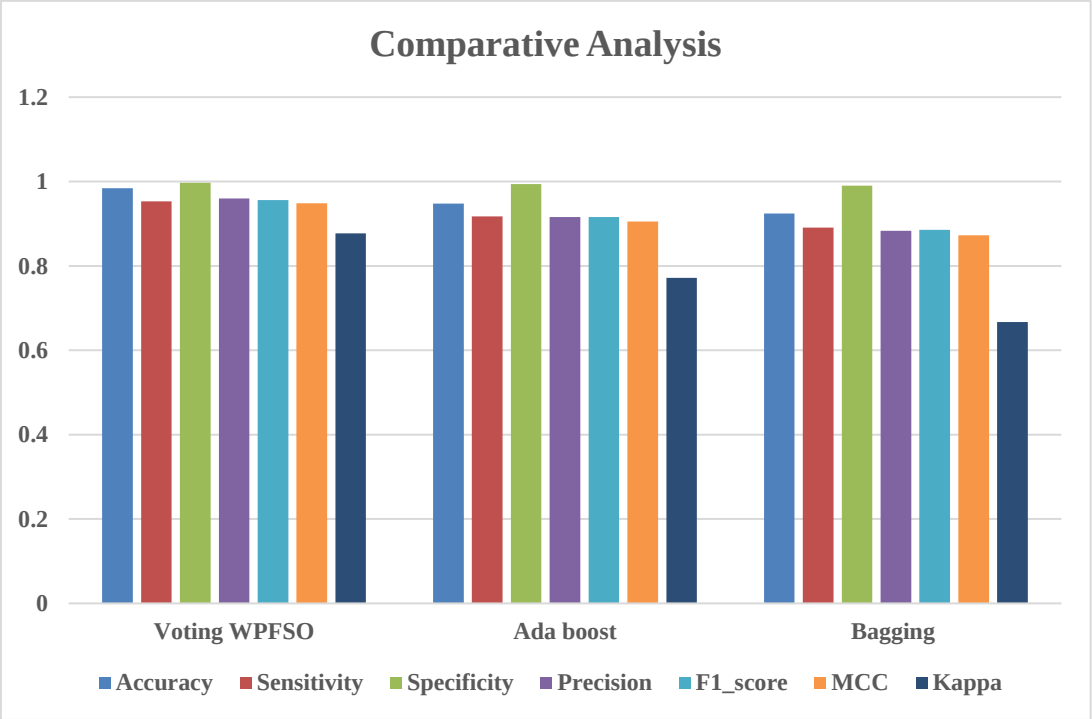


Figure 9 Comparative analysis of the proposed and existing voting classifier

5. Conclusion:

This paper affords a unique Intrusion Detection gadget (IDS) is presented for enhancing community overall performance through the correct detection of disbursed Denial-of-provider (DDoS) assaults in cloud surroundings. The proposed IDS framework is ready with a novel genetic algorithm based arithmetic optimization set of rules for attack vector selection and an wise ensemble voting Classifier primarily based on Woodpecker based totally Flamingo seek optimization set of rules for DDOS attack detection and category. The proposed IDS outperforms diverse present methods in terms of community performance, DDOS attack prediction and mitigation fee. The accuracy charge of assault detection in proposed machine is ninety 8.4% which is comparably best than present methods.

Doi:10.48047/ijiee.2025.15.5.32

6. References

[1] J. He, “Cloud computing load balancing mechanism taking into account load balancing ant colony optimization algorithm,” Computational Intelligence and Neuroscience, vol. 2022, Article ID 3120883, 10 pages, 2022.

[2] F. S. Al-Anzi, S. K. Yadav, and J. Soni, “Cloud computing: security model comprising governance, risk management and compliance,” in Proceedings of the 2014 International Conference on Data Mining and Intelligent Computing (ICDMIC), September 2014.

[3] N. Mishra, R. K. Singh, and S. K. Yadav, “Design a new protocol for vulnerability detection in cloud computing security improvement,” in Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2021, Delhi, India, February 2021.

[4] D. Alsmadi and V. Prybutok, “Sharing and storage behavior via cloud computing: security and privacy in research and practice,” Computers in Human Behavior, vol. 85, pp. 218– 226, 2018.

[5] J. Shroff, R. Walambe, S. K. Singh, and K. Kotecha, “Enhanced security against volumetric DDoS attacks using adversarial machine learning,” Wireless Communications and Mobile Computing, vol. 2022, Article ID 5757164, 10 pages, 2022.

[6] O. Cepheli, S. B¨uy¨ukçorak, and G. Karabulut Kurt, “Hybrid ¨ intrusion detection system for ddos attacks,” Journal of Electrical and Computer Engineering, pp. 1–8, 2016.

[7] B. A. Khalaf, S. A. Mostafa, A. Mustapha et al., “An adaptive protection of flooding attacks model for complex network environments,” Security and Communication Networks, vol. 2021, Article ID 5542919, 17 pages, 2021.

[8] M. Myint Oo, S. Kamolphiwong, T. Kamolphiwong, and S. Vasupongayya, “Advanced support vector machine- (ASVM-) based detection for distributed denial of service (DDoS) attack on software defined networking (SDN),” Journal of Computer Networks and Communications, Article ID 8012568, 12 pages, 2019.

Doi:10.48047/ijiee.2025.15.5.32

[9] X. Yu, W. Yu, S. Li, X. Yang, Y. Chen, and H. Lu, "WEB DDoS attack detection method based on semisupervised learning," Security and Communication Networks, vol. 2021, Article ID 9534016, 10 pages, 2021.

[10] U. A. Butt, M. Mehmood, S. B. H. Shah et al., "A review of machine learning algorithms for cloud computing security," Electronics, vol. 9, no. 9, p. 1379, 2020.

[11] M. Idhammad, K. Afdel, and M. Belouch, "Detection system of HTTP DDoS attacks in a cloud environment based on information theoretic entropy and random forest," Security and Communication Networks, vol. 2018, Article ID 1263123, 13 pages, 2018.

[12] M. Eshtay, H. Faris, and N. Obeid, "A competitive swarm optimizer with hybrid encoding for simultaneously optimizing the weights and structure of Extreme Learning Machines for classification problems," International Journal of Machine Learning and Cybernetics, vol. 11, no. 8, pp. 1801– 1823, 2020.

[13] N. Mishra and R. K. Singh, "DDoS vulnerabilities analysis and mitigation model in cloud computing," Journal of Discrete Mathematical Sciences and Cryptography, vol. 23, no. 2, pp. 535–545, 2020.

[14] C. Amine, S. Ben Alla, and A. Ezzati, "Makespan optimisation in cloudlet scheduling with improved DQN algorithm in cloud computing," Scientific Programming, vol. 2021, Article ID 7216795, 11 pages, 2021.

[15] A. Amjad, T. Alyas, U. Farooq, and M. Tariq, "Detection and mitigation of DDoS attack in cloud computing using machine learning algorithm," ICST Transactions on Scalable Information Systems, vol. 0, no. 0, Article ID 159834, 2018.

[16] A. Aliyu, A. H. Abdullah, O. Kaiwartya et al., "Mobile cloud computing: taxonomy and challenges," Journal of Computer Networks and Communications, Article ID 2547921, 23 pages, 2020.

[17] P. Singh and V. Ranga, "Attack and intrusion detection in cloud computing using an ensemble learning approach," International Journal of Information Technology, vol. 13, no. 2, pp. 565–571, 2021.

[18] H. A. Mahmood and S. H. Hashem, "Network intrusion detection system (NIDS) in cloud environment based on hidden na¨ıve bayes multiclass classifier," Al-Mustansiriyah Journal of Science, vol. 28, no. 2, pp. 134–142, 2018.

[19] Y. Jiang, B. Qiu, C. Xu, and C. Li, "&e research of clinical decision support system based on three-layer knowledge base model," Journal of healthcare engineering, Article ID 6535286, 8 pages, 2017. [20] Y. Berguig, I. Laassiri, and S. Hanaoui, "DoS detection based on mobile agent and na¨ıve bayes filter," in Proceedings of the 2018

Doi:10.48047/ijiee.2025.15.5.32

International Symposium on Advanced Electrical and Communication Technologies (ISAECT), Rabat, Morocco, November 2018.

[20] S. Nandi, S. Phadikar, and K. Majumder, "Detection of DDoS attack and classification using a hybrid approach," in Proceedings of the 2020 >ird ISEA Conference on Security and Privacy (ISEA-ISAP) 2020, pp. 41–47, Guwahati, India, March 2020.

[21] J. Kim, J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," Electronics, vol. 9, no. 6, p. 916, 2020.

[22] A. E. Cil, K. Yildiz, and A. Buldu, "Detection of DDoS attacks with feed-forward based deep neural network model," Expert Systems with Applications, vol. 169, no. December 2020, Article ID 114520, 2021.

[23] A. Rangapur, T. Kanakam, and A. Jubilson, "DDoSDet: an approach to Detect DDoS attacks using Neural Networks," 2022, https://arxiv.org/abs/2201.09514.

[24] A. Saroha and A. Singh, "Security concerns and countermeasures in cloud computing: a qualitative analysis," International Journal of Information Technology, vol. 11.4, pp. 683–690, 2019.

[25] R. Goel, M. Garuba, and A. Girma, "Cloud computing vulnerability: DDoS as its main security threat, and analysis of IDS as a solution model," in Proceedings of the 2014 11th International Conference on Information Technology: New Generations, Las Vegas, NV, USA, April 2014.

[26] R. V. Deshmukh and K. K. Devadkar, "Understanding DDoS attack & its effect in cloud environment," Procedia Computer Science, vol. 49, pp. 202–210, 2015.

[27] M. Masdari and M. Jalali, "A survey and taxonomy of dos attacks in cloud computing," Security and Communication Networks, vol. 9, 2016.

[28] P. Oberoi, "Survey of various security attacks in clouds based environments," International Journal of Advanced Research in Computer Science, vol. 8, no. 9, pp. 405–410, 2017.

[29] E. S. JeyaJothi, J. Anitha, S. Rani, and B. Tiwari, "A comprehensive review: computational models for obstructive sleep apnea detection in biomedical applications," BioMed Research International, pp. 1–21, 2022.

[30] N. O. Ogwara, K. Petrova, and M. L. Yang, "Towards the development of a cloud computing intrusion detection framework using an ensemble hybrid feature selection approach," Journal of Computer Networks and Communications, vol. 2022, Article ID 5988567, 16 pages, 2022.

[31] UNB, "NSL-KDD dataset," https://www.unb.ca/cic/datasets/ nsl.html.

Doi:10.48047/ijiee.2025.15.5.32

[32] Kaggle, "KDD-CUP-98 Dataset," https://www.kaggle.com/datasets/towhidultonmoy/kddcup98-dataset.

[33] Kaggle, "NSL-KDD randomforest w/optuna,".

[34] L. Dhanabal and S. P. Shantharajah, "A study on NSL-KDD dataset for intrusion detection system based on classification algorithms," International journal of advanced research in computer and communication engineering, vol. 4.6, pp. 446– 452, 2015.

[35] M. Li and D. Han, X. Yin, H. Liu, and D. Li, Design and implementation of an anomaly network traffic detection model integrating temporal and spatial features," Security and Communication Networks, vol. 2021, Article ID 7045823, 15 pages, 2021.

[36] S. K. Yadav, D. K. Tayal, S. N. Shivhare, and S. Naresh Shivhare, "Perplexed Bayes classifier-based secure and intelligent approach for aspect level sentiment analysis," International Journal of Advanced Intelligence Paradigms, vol. 13, no. 1/2, p. 15, 2019.

[37] A. Ul Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," Mobile Information Systems, vol. 2018, Article ID 3860146, 21 pages, 2018.

[38] C. S. Carlos, "Perplexed Bayes classifier," in Proceedings of the 12th International Conference on Natural Language Processing, Trivandrum, India, December 2015.

[39] K. Dhingra and S. K. Yadav, "Spam analysis of big reviews dataset using fuzzy ranking evaluation algorithm and hadoop," International Journal of Machine Learning and Cybernetics, vol. 10.8, pp. 2143–2162, 2019.

Doi:10.48047/ijiee.2025.15.5.32