International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

DDoS Attacks Detection via Three-Tier Optimization Algorithm in Cloud Environment

[1]M. Mohan*, [2]Dr. V. Tamizhazhagan, [3]Dr. S. Balaji

Submitted: 24/12/2023 **Revised**: 04/02/2024 **Accepted**: 13/02/2024

Abstract: Cloud environment is significantly threatened by distributed denial of service (DDoS) attacks, which can quickly reduce the system's resources, keep the servers busy, and cause severe system damage. Recently, many DDoS attacks have been carried out using cunning tactics, including low-rate attacks and attacks that target authenticated users. In this paper, a novel DDoS Attack Detection (DAD) technique has been proposed to detect and recover the attacker. The suggested technique involves preprocessing process in terms of three processes including Sequential Backward Selection (SBS), Independent Component Analysis (ICA), and Sewing Training-Based Optimization (STBO). An SBS analyzes and detects unauthorized or authorized requests, as well as recover attacks. But some attacks will not recover in beginning stage so the further preprocessing process will perform using ICA under identify, retrieval, organizer. The Identification process will check whether the request is from attacker or not from the database using SBS algorithm. If the request from unauthorized person, the retrieval process will start to recover and after recover the request to store data into the organizer. Finally, the performance of the suggested DAD technique is compared with MLDMF, PSD, and ADADM and the performance analysis of the proposed technique is determined using detection rate, precision, False alarm rate, and accuracy. The proposed method achieves a higher level of accuracy is 97%.

Keywords: *Cloud environment, Sequential Backward Selection, Independent Component Analysis, Swarm Intelligence, unauthorized person.*

1.Introduction

Cloud computing is a commonly used technique that performs complex computing and massive-scale, which is large number of storages, services, and information's are accessible over the Internet [1]. The cloud environment has to cope with a number of security issues because cloud services are distributed by nature and might be shared by millions of users; DDoS attack is almost one among the critical [2,3]. The scattered and open structure of the cloud and the lack of user control over its resources make it an appealing target for attackers [4].

DDoS attack is occurring from various locations to attack a single victim since cloud computing exhibits the features of accessibility, adaptability, availability, and scalability [5]. Most hosts are taking part in an attack. The attacker wizard launches the DDOS attack by managing agents with two or more processing controllers [6].

DDoS attacks are causing financial damage to companies and websites all around the world since they are continually growing and are intended to disrupt major corporations [7]. In some situations, attackers claim to hurting a unit or a business that is intended to launch a traffic; in other instances, the intruder merely tries to violate the target and cause the most severe injury or issue. Once a DoS attack has been determined, an attacker target may appear. [8].

In reality, lot of large-scale DoS attacks are spread, which is launched from several computers [9]. One source of DDoS attacks can easily reduce the severity of the attacks because the defenders can block the network traffic from attacker sources, but it is challenging to recognise and defend against attacks on different attack systems. So, it is very difficult to detect the difference malicious packets and legitimate internet traffic [10]. In order to find detect and recover the attack, the machine learning DAD technique has been proposed.

The major contributions of proposed DAD technique are given below:

- The proposed technique used SBS, ICA, and STBO algorithm for detect whether the request is known or unknown person in the database. If the request from unauthorized person, the retrieval process will happen.
- Received attack and recover attack data can store into the database for future reference.
- The performance analysis of the proposed technique is determined using detection rate, precision, False alarm rate, and accuracy.

The remaining portions of the analysis are structured as follows: Section II describe the Literature survey in detail.

[1] *Research Scholar, Department of Information Technology, Annamalai University. Tamil Nadu, India,*
[2] *Research Supervisor , Assistant Professor, Department of Information Technology , Annamalai University, Tamil Nadu, India.*
[3] *Professor, CSE, Panimalar Engineering College, Chennai, Tamilnadu, India.*
** Corresponding Author Email: mohan.rm@gmail.com*

Section III describes the suggested DAD technique. The result is given in section IV and finally section V describe the conclusion.

2.Literature survey

There are various DDoS detection techniques that are based on various architectures, including source-end, in-network, and victim-end. These techniques contain knowledge-based, data mining, statistical, soft computing, and machine learning techniques. An overview of a few recent advancement and techniques for security management within this part.

In 2018, Yan, Q., et al., [11] proposed a multi-level DDoS mitigation framework (MLDMF) for Industrial Internet of Things (IIoT), which comprises the fog computing level, edge computing level, and cloud computing level. In the IIoT, DDoS assaults are reduced, and a huge number of IIoT devices are managed using software-defined networking. The usefulness of the suggested framework has been demonstrated through experiments.

In 2018, Idhammad, M., et al., [12] proposed a HTTP DDoS attacks in the cloud based on information-theoretic entropy and the Random Forest ensemble learning algorithm. On the publicly available CIDDS-001 dataset, a number of experiments were run to evaluate the suggested technique. With a running time of 18.5 seconds, an accuracy of 99.54%, and an FPR of 0.4%, the suggested method produces satisfactory results.

In 2020, Virupakshar, K.B., et al., [13] suggested a cloud operating system with a built-in firewall, DDoS detection software, an Open Stack integrated firewall, and raw socket programming for network traffic monitoring. The dataset DNN, KNN, and Decision Tree algorithms used in DDoS attacks. DDoS attacks are discovered as a result of the experiment, and the private cloud administrator is contacted.

In 2020, Gumaste, S. and Shinde, [14] proposed employing machine learning classifiers on a distributed processing infrastructure to detect DDoS attacks in real time. It assesses the DDoS detection system on a cloud built on Open Stack and tested with the Apache Spark framework. The experimental results demonstrate that the random forest method provides a more accurate classifier.

In 2020, Saxena, R. and Dey, S., [15] proposed a packet traceback system based on a third-party auditor (TPA), which is used to determine where the DDoS attack originated. Because of its powerful identification element, which is dependent on the vulnerabilities the intruder left behind, it offers an effective and fruitful solution. In order to generate attack alerts for various cloud users, it analyses the traffic patterns. The experimental findings are presented to demonstrate the efficacy of the suggested approach for DDoS attack mitigation and prevention.

In 2021, He, Q., et al., [16] proposed two strategies for preventing edge DDoS attacks: an ideal strategy called EDMOpti and a game-theoretic strategy called EDMGame. In order to locate the Nash equilibrium and solve the EDM problem, EDMGame formulates and the EDM problem as a possible EDM game that accepts a Nash equilibrium. The experimental finding demonstrates how to successfully and efficiently tackle the EDM problem.

These methods outperform those previously developed but they have some drawbacks such as a loss of data, privacy violations, high service cost, hosting, and server problems. To overcome the above drawbacks, a novel DAD technique has been suggested in this paper.

3.Proposed method

In this section, a novel DAD technique is used to recover the data from DDoS attack. Figure 1 illustrate pre-processing process in terms of three process including SBS, ICA and STBO. SBS is used to analyse and detect abnormal or authorised requests, as well as to recover attacks. However, some attacks will not recover in the early stages, so the subsequent pre-processing process will use ICA under identification, retrieval, and organisation. Using the SBS algorithm, the identification process will determine whether the request is for a known or unknown person in the database. If the request is made by an attacker, the STBO optimizer will begin the retrieval process, and then it will be saved in the organizer.

3.1 Pre processing

The first step is to pre-process the DDoS attacked data. To recover the traffic using pre-processing techniques such as ICA.

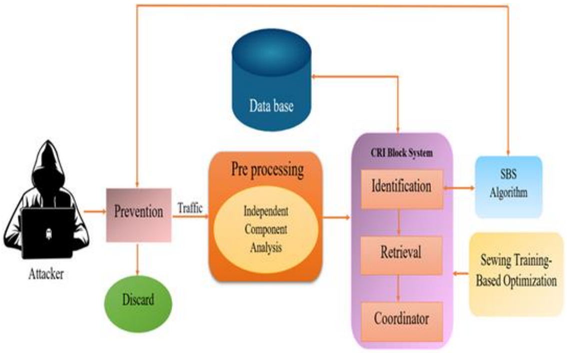

Fig 1: Overall block diagram of proposed DAD system

3.1.1 Independent Component Analysis

The statistical independence property is the foundation of the multivariate signal-processing approach known as ICA. A series of observations composed of linear mixtures of the underlying sources are subjected to ICA algorithms in an effort to isolate signals from independent sources. The information is projected onto some new axes, the origins of which must be identified. The term "blind source separation" refers as set of algorithms that separate mixed signals into original sources as a result of this approach. An ideal illustration of isolating a mixed signal is gathering with a live band. The cocktail party is listening to a synthesis of instruments, voices, and background noises rather than to each band member's instrument individually.

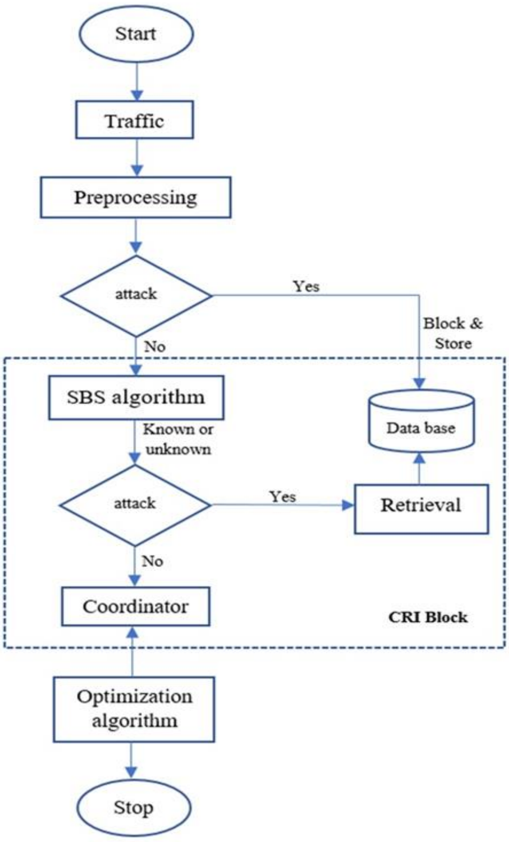

Fig 2: Flow diagram of proposed DAD system

Let's take a look at M time series, each with N samples. Finding a way to change these time series into a new representation that distinguishes and separates independent components is the goal.

Formally, it is equivalent to the M measured time series.

$$y_i = (y_{i1}, y_{i2}, \dots \dots \dots, y_{iN})^T \text{ , } i = 1, \dots \dots . ., M \qquad (1)$$

Matrix Y rows are the transposed into time series

$$Y = \begin{pmatrix} y_{11} & \cdots & y_{1N} \\ \vdots & \ddots & \vdots \\ y_{M1} & \cdots & y_{MN} \end{pmatrix} \qquad (2)$$

This M×N matrix is intended to the linear combination of the original signals, which can alternatively be represented by another M×N matrix and P with a structure identical to Y, provided that the rows of P are the transposed relative to the original time series

$$P_i = (P_{i1},\ P_{i2}, \dots \dots, P_{iN})$$

The linear combination can be stated as follows:

$$Y = AP, \qquad (3)$$

Finding the mixing matrix, A and original sources P is the primary goal of ICA. The problem formulation for this activity of inverse dual problem. Finding a demixing matrix Z is necessary before the source vector can be determined using this matrix."

$$P = ZY, \qquad (4)$$

It is impossible to calculate Z or A directly. Instead, a demixing matrix Z is calculated that works on Y to provide the estimate

X ≈ P ˆof the sources.

$$X = ZY = P\hat{}\text{and } Z \approx\ A^{-1} \qquad (5)$$

The ICA approach uses high-order statistics to measure the signal-to-noise separation and some factorization of the observed data to approximate the signal. Because of its non-Gaussian character, most ICA algorithms inverse of A to calculate Z, which enables the source vector to be determined. Finding the A^(-1) that optimises the non-Gaussian nature of the independent components is the key underlying this technique. Usually, maximum-likelihood estimation, output entropy maximisation, or output mutual information minimization are used to carry out this operation.

The fastICA negentropy notion is used to quantify the non-Gaussian character. The formula for calculating negentropy K is as follows:

$$K(y)=H(y_{gauss})–H(y) \qquad (6)$$

The fastICA algorithm is based on a fixed-point method for finding Z ≈ A^(-1) by maximization the negentropy. In addition, some attack will not clear in beginning stage so the subsequent pre-processing process will use ICA under identification, retrieval, and organizer using SBS and SI Algorithm.

3.2 CRI Block

Some traffic will not recover using ICA technique. In this paper, suggested the CRI block with various process such as identification, retrieval, and coordinator as shown in figure 1. These processes are described below:

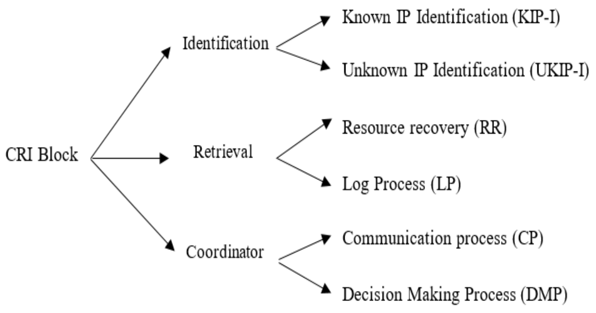
CRI Block
Identification
Known IP Identification (KIP-I)
Unknown IP Identification (UKIP-I)
Retrieval
Resource recovery (RR)
Log Process (LP)
Coordinator
Communication process (CP)
Decision Making Process (DMP)

Fig 3: Classification of CRI Block

3.2.1 Identification

The identification process, which uses the SBS algorithm to search for various DDoS attacks, will be active for the majority of the time. It uses two different processes such as, Known IP Identification (KIP-I), Unknown IP (UKIP-I) for identify the IP address of incoming request.

3.2.1.1 Known IP Identification (KIP-I)

The source IP of the receiving request is used to determine the attack's identity. An attack may occasionally be launched by the current user using a known IP. Because it would be extremely costly to miss this type of conduct, this agent has been put in place. If confirm the IP address is known person automatically start the recovery process.

3.2.1.2 Unknown IP Identification (UKIP-1)

The new IP address is only responsible for the majority of attacks. These attacks will directly be affected to the bandwidth and some technology deployed to detect these kinds of attacks. As a result, this process will concentrate on checking new incoming IP addresses for suspicious activity. Finally, store all the details into the database to retrieve the traffic.

3.2.2 Retrieval

This process is triggered by the identification block when the attacker's identification is confirmed start the recovery process automatically. It done by using two process such as, Resource recovery (RR), and Log process (LP).

3.2.2.1 Resource recovery (RR)

This agent will use the cloud resources to allocate them to the specific user identified as the intruder by using SBS algorithm.

3.2.2.2 Log process (LP)

All the data already stored in the database. It will maintain track of recent attacks as well as the source IP address and behaviour of the attackers, enabling the DMP to make better-informed decisions.

3.2.3 Coordinator

This process is interconnected with all processes using STBO algorithm to take correct decision. It uses two different process such as, Communication process (CA), and Decision-making process (DMP).

3.2.3.1 Communication process (CA)

In order to ensure, communication is safe against internal or internal attacks, it serves as the backbone of communication among the other agents. Information flow is encrypted through this process.

3.2.3.2 Decision making process (DMP)

The DMP contains many steps. Any of the identifications can start a DDoS attack detection, but intelligence is required to confirm those attacks, and these DMPs provide that information. These processes maintain this security system as an autonomously improving one.

3.3Sequential Backward Selection Algorithm

It identifies whether the request is known or unknown person in the database by using SBS algorithm. After identify the attacker recover process will happen and coordinate all the data into the database.

SBS is a traditional feature selection algorithm that minimises execution time and discovers feature space and subspace features with the least amount of latency. In some circumstances, SBS can help the model's predictive power. For instance, SBS can delete features from the entire feature space if a method is overfitting until a new feature subspace is generated. The criterion, which is just the difference in classifier performance before and after the exclusion of a particular feature, is computed by the criterion to identify. It should be deleted from the feature space at each step necessary to construct a function of criterion J to minimize. The trait that maximizes the criterion is the one that is removed at each stage.

Pseudo-code for SBS algorithm's

Input: preprocessed request

Output: detection rate

```
Step 1: Initialize s = g, where s is the dimension of the feature full space Y_g .
Step 2: Remove the criterion-maximizing feature y-, y- = arg max J (Y_s - y),
Where y ∈ Y_s
Step 3: Remove feature y- from the feature space as follows:
 Y_s-1≔ Y_s-y- ;s≔s-1.
Step 4: Finish if s has reached the required features; otherwise, repeat Step 2.
```

3.4 Sewing Training Based Optimization Algorithm

The STBO algorithm is based on metaheuristic algorithm made up of inexperienced tailors and trainers. The recommended values for the decision variables are represented by STBO members as a potential solution to the problem. As a result, each member of the STBO may be mathematically represented by a vector, while the STBO population can be represented by a matrix.

$$\mathrm{Y}=\begin{bmatrix} Y_1 \\ \vdots \\ Y_j \\ \vdots \\ Y_M \end{bmatrix}_{M\times n} = \begin{bmatrix} y_{1,1} & \dots & y_{1,i} & \cdots & y_{1,n} \\ \vdots & \ddots & \vdots & \iddots & \vdots \\ y_{j,1} & \dots & y_{j,i} & \dots & y_{j,n} \\ \vdots & \iddots & \vdots & \ddots & \vdots \\ y_{M,1} & \cdots & y_{M,i} & \cdots & y_{M,n} \end{bmatrix}_{M\times n} \qquad (7)$$

where Y represents the STBO matrix, Y_j represents the jth STBO's member, M represents the number of STBO population members, and n represents the number of problem variables. All population members are randomly initialised using an equation 8 at the start of the STBO implementation.

$$y_{j,i} = ub_i + s.(lb_i - ub_i), j = 1,2,\dots,M\ , i = 1,2,\dots,n,$$
(8)

where y_(j,i) is an ith variable determined by the jth STBO's member y_j, r is a random number of the range [0,1], 〖ub〗_i and 〖lb〗_i are the lower and upper bound of the jth variable.

Each STBO participant represents the prospective fix issue at hand. As a result, using the values provided by each potential solution, the objective function of the issue may be assessed. Based on the placement of potential solutions in the problem variables, a vector equation 9 can be used to model the values determined for the goal function.

$$\mathrm{G}=\begin{bmatrix} G_1 \\ \vdots \\ G_j \\ \vdots \\ G_M \end{bmatrix}_{M\times 1} = \begin{bmatrix} G(Y_1) \\ \vdots \\ G(Y_j) \\ \vdots \\ G(Y_M) \end{bmatrix}_{M\times 1} \qquad (9)$$

where Fi is the objective function value for the ith candidate solution and G is the objective function vector.

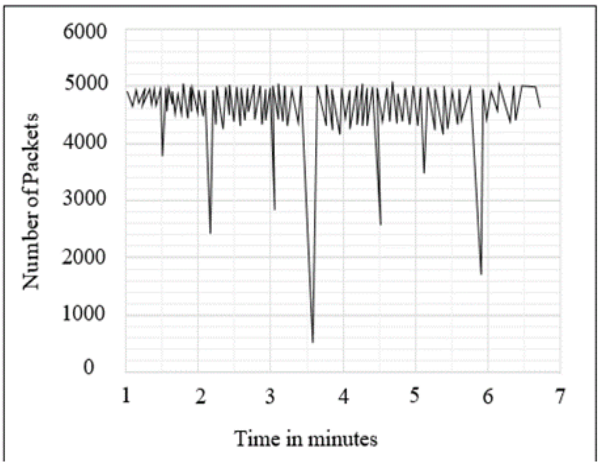

Fig 4 (a): Constant rate attack profile

Results and Disscussion

The proposed DAD technique acts as an essential role in evaluating the DDoS attack detection performance of tolerance factor. This tolerance factor helps for the prediction of the threshold range.

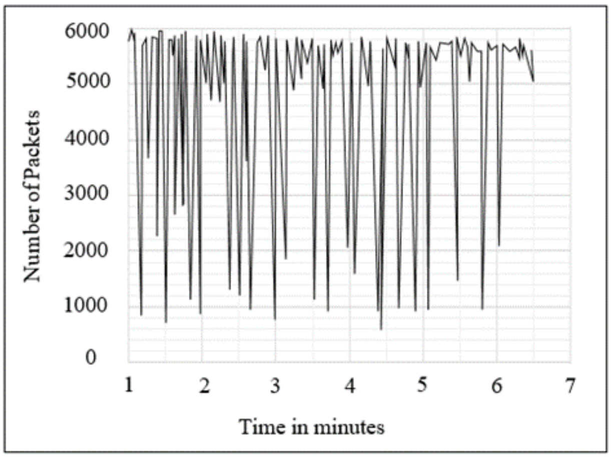

Fig 4 (b): Pulsating attack

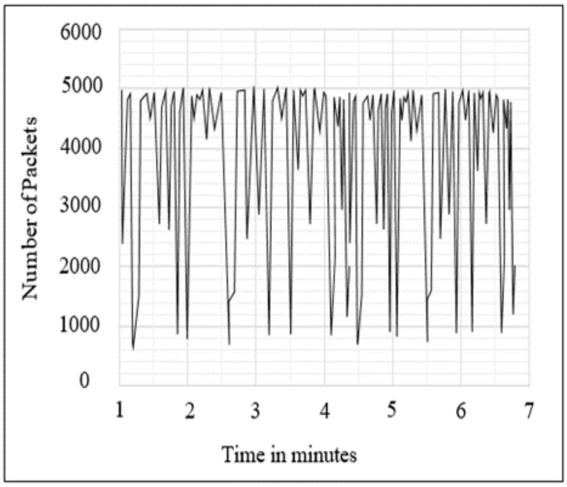

Fig 4 (c): Subgroup attack profile

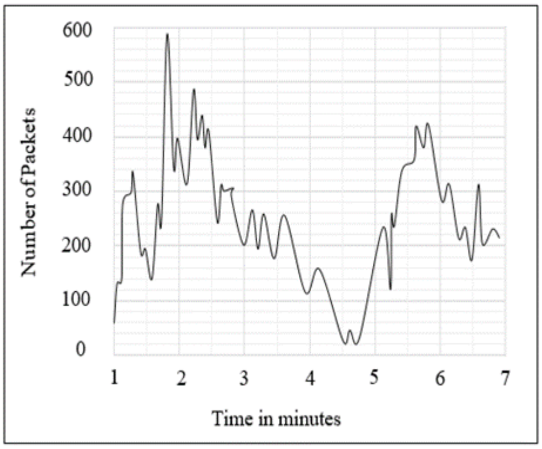

Fig 4 (d): Increasing rate attack profile

Figure 4(a), 4(b), 4(c)and 4(d) shows different DDoS attack such as Constant rate, Pulsing rate, Subgroup and Increasing rate. x axis represents time in minutes and y axis represents number of packets. A proposed DAD technique is designed to handle the standard profile and the attack profile. The deviation of more significant to the perfect threshold value. It takes the threshold value in between the range 0.025 to 0.990 for CIC-DDoS2019. Increasing rate attack is a Low-Rate DDoS attack because it achieves an ongoing attack.

4.1 Comparative Analysis

There is a comparison between traditional methods and the proposed model in this division as well. Based on a comparison of performance with existing approaches, this technique is more efficient.

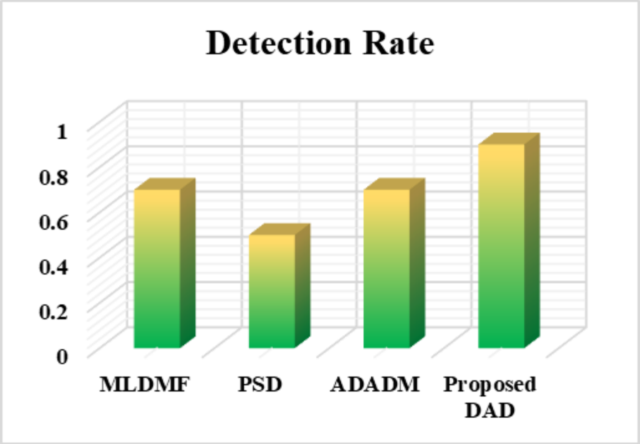

Fig 5(a): Detection rate comparison

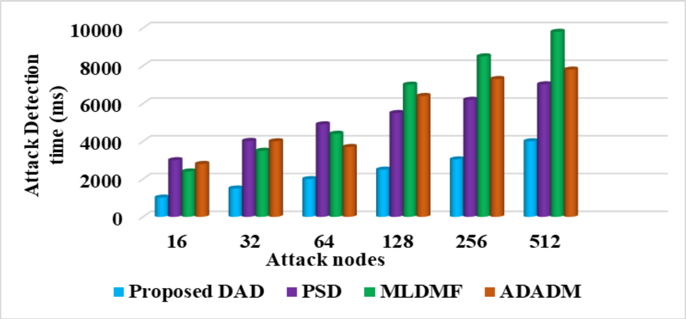

Fig 5(b): Precision range comparison

Figure 5(a) illustrates the detection rate of proposed DAD technique. It compared to MLDMF, PSD, and ADADM the proposed DAD technique shows the detection rate and reaches around 80 % at a tolerance rate of 0.8 although all three methods started at around 89%. Figure 5(b) shows that DAD technique keeps getting better until it reaches 90%, while MLDMF, PSD, and ADADM limit at a tolerance factor of 0.97

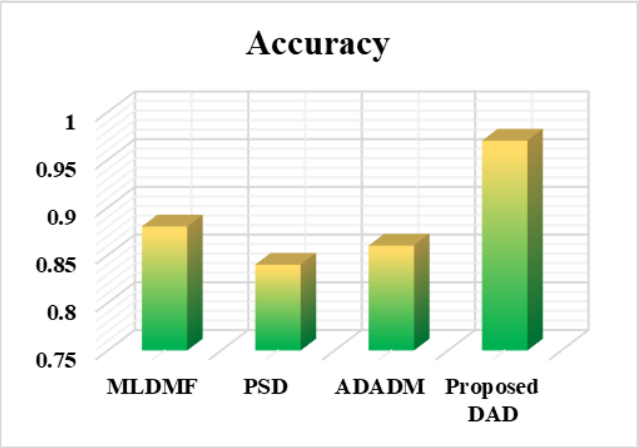

Fig 5(c): False alarm rate comparison

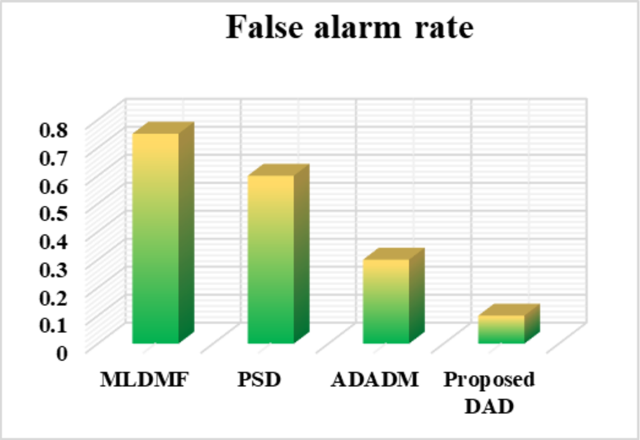

Fig 5(d): Accuracy comparison

Figure 5(c) depicts the increasing tolerance factor and shows that our DAD detection technique outperforms MLDMF, PSD, and ADADM by a wide margin. In the 0.97 threshold limit, MLDMF, PSD, and ADADM score 88%, 84%, and 86% accuracy, respectively, in figure 5(d). When these three are compared, the ADAPT metric achieves a better result for the detection of signature attacks on the cloud platform.

From table 1 compares the various algorithms such as Naive Bayes, Nearest Neighbors, Random Forest, AdaBoost, Neural Net, Linear SVM, ICA, SBS and it determined the DDOS attack at the highest rate and in the shortest time. It would be instructive to also list in Table 1 the corresponding optimal scores by the other machine learning algorithms. And it clearly indicates that the proposed SBS achieved better result than other algorithms.

Algorithm	Success Rate (%)	Detection Time (sn)
Naive Bayes	68,8	694
Nearest Neighbors	69,6	675
Random Forest	70,3	700
AdaBoost	70,7	706
Neural Net	71,6	719
Linear SVM	72,6	727
ICA	84,2	784
Our proposed (SBS)	91,9	893

Attack detection time is the total amount of time needed to identify the attacker's IP address, identify susceptible behaviour, and contact the coordinator to validate other properties. Fig. 6 displays the average amount of time needed to identify a single attack. When compared to existing approaches, the suggested DAD has the quickest attack detection time and is about 65% more effective than the PSD, MLDMF, and ADADM.

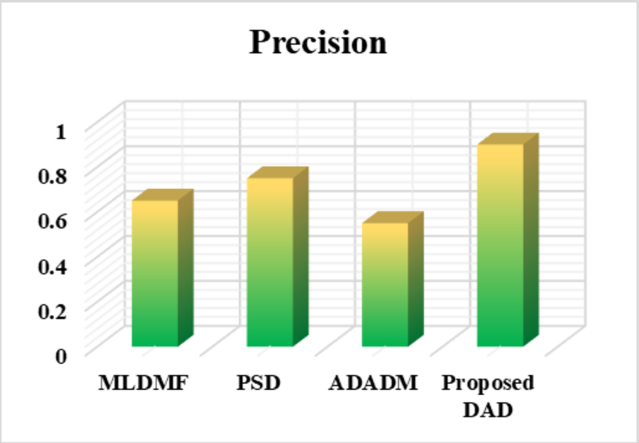

Fig 6: Attack detection time

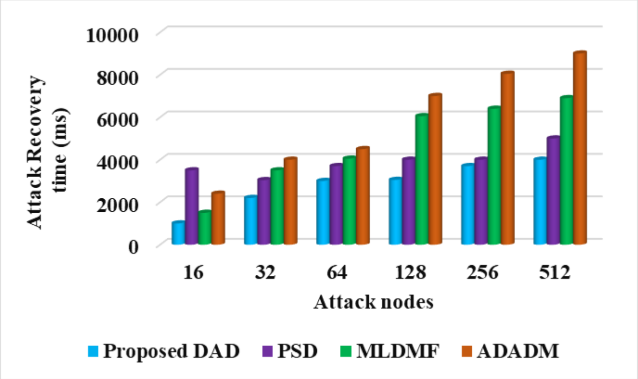

Fig 7: Attack recovery time

Attack recovery time is the amount of time needed to release the resources allotted to the IP address of the identified attacker after verification with DMP. This process entails two steps: resource recovery and updating the log details. Fig. 7 displays the typical amount of time needed to recover after a single attack. The proposed DAD gives an accuracy of 97% better than the PSD, MLDMF, and ADADM.

5. Conclusion

In this paper, a novel DDoS Attack Detection (DAD) technique has been proposed to detect and recovery the attacker. The suggested technique involves preprocessing process in terms of three processes including Sequential Backward Selection (SBS), Independent Component Analysis (ICA), and Sewing Training-Based Optimization (STBO). An SBS analyzes and detects unauthorized or authorized requests, as well as recover attacks. But some attacks will not recover in beginning stage so the further preprocessing process will perform using ICA under identify, retrieval, organizer. The Identification process will check whether the request is from attacker or not from the database using SBS algorithm. If the request from unauthorized person, the retrieval process will start to recover and after recover the request to store data into the organizer. Finally, the performance of the suggested DAD technique is compared with MLDMF, PSD, and ADADM and the performance analysis of the proposed technique is determined using detection rate, precision, False alarm rate, and accuracy. The proposed method achieves a higher level of accuracy is 97%. This work can be extended in the future to use stochastic-based filtering for more optimization in the attack detection phase.

6.References

[1] Mohaiminul Islam, Shamim Reza. The Rise of Big Data and Cloud Computing, Internet of Things and Cloud Computing. Volume 7, Issue 2, June 2019 , pp. 45-53.

[2] Aytaç T, Aydın MA, Zaim AH. Detection DDOS Attacks Using Machine Learning Methods. Electrica, 2020; 20(2): 159-167.

[3] Santos, R., Souza, D., Santo, W., Ribeiro, A. and Moreno, E., 2020. Machine learning algorithms to detect DDoS attacks in SDN. Concurrency and Computation: Practice and Experience, 32(16), p.e5402.

[4] Mohammed, M.H., 2022. An enhancement of cyber security management for opportunistic systems. Measurement: Sensors, 24, p.100547.

[5] Hesham Abusaimeh, “Distributed Denial of Service Attacks in Cloud Computing” International Journal of Advanced Computer Science and Applications(IJACSA), 11(6), 2020.

[6] Mohammad, M.A. and Jawhar, M.M., 2022. Compare between PSO and artificial bee colony

optimization algorithm in detecting DoS attacks from network traffic. TELKOMNIKA Telecommunication Computing Electronics and ControlVol. 20, No. 4, August 2022, pp. 780~787.

[7] Roopak, M., Tian, G. Y., Chambers, J. A. (2020). Multi-objective-based Feature Selection For Ddos Attack Detection In Iot Networks. IET Networks, 3(9), 120-127

[8] Virupakshar, K.B., Asundi, M., Channal, K., Shettar, P., Patil, S. and Narayan, D.G., 2020. Distributed denial of service (DDoS) attacks detection system for OpenStack-based private cloud. Procedia Computer Science, 167, pp.2297-2307

[9] AM R. A DDoS defence framework in software defined network using ensemble classifier with rough set theory based feature selection. International Journal of Advanced Technology and Engineering Exploration. 2021; 8(82):1120-1135.

[10] Khraisat, A., Gondal, I., Vamplew, P., Kamruzzaman, J. and Alazab, A., 2019. A novel ensemble of hybrid intrusion detection system for detecting internet of things attacks. Electronics, 8(11), p.1210.

[11] Yan, Q., Huang, W., Luo, X., Gong, Q. and Yu, F.R., 2018. A multi-level DDoS mitigation framework for the industrial Internet of Things. IEEE Communications Magazine, 56(2), pp.30-36.

[12] Mohamed Idhammad, Karim Afdel, Mustapha 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.

[13] Virupakshar, K.B., Asundi, M., Channal, K., Shettar, P., Patil, S. and Narayan, D.G., 2020. Distributed denial of service (DDoS) attacks detection system for OpenStack-based private cloud. Procedia Computer Science, 167, pp.2297-2307.

[14] Gumaste, S. and Shinde, S., 2020. Detection of DDOS attacks in openstack-based private cloud using apache spark. Journal of Telecommunications and Information Technology, 62-71.

[15] Saxena, R. and Dey, S., 2020. DDoS attack prevention using collaborative approach for cloud computing. Cluster Computing, 23(2), pp.1329-1344.

[16] Q. He, et al.,"A Game-Theoretical Approach for Mitigating Edge DDoS Attack" in IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 04, pp. 2333-2348, 2022.

[17] Yadav, B. and Satija, R., Introducing three-tier captcha to prevent DDOS attack in Cloud computing. International Journal on Recent and Innovation Trends in Computing and Communication, 2(7), pp.1969-1973.

[18] Hou, F., Sun, J., Yang, Q. and Pang, Z., 2022. Deep reinforcement learning for optimal denial-of-service attacks scheduling. Science China Information Sciences, 65(6), pp.1-9

[19] Yousuf, O. and Mir, R.N., 2022. DDoS attack detection in Internet of Things using recurrent neural network. Computers and Electrical Engineering, 101, p.108034.

[20] Adi, E., Anwar, A., Baig, Z. and Zeadally, S., 2020. Machine learning and data analytics for the IoT. Neural Computing and Applications, 32(20), pp.16205-16233.