Detection of Distributed Denial of Service Attack in Cloud Environment

Subhash Kumar*, Dr. Rakesh Kumar**
*Department of Computer Science & Engineering
Rabindranath Tagore University, India
Email- bhartikr14032@gmail.com
**Department of Computer Science & Engineering
Rabindranath Tagore University, Raisen, Madhya Pradesh, India
Email- rakeshmittan@gmail.com

Abstract-
Within the realm of cloud computing, a significant concern pertains to intrusion detection and mitigation, given its potential to disrupt the entire architecture's functionality. Countless efforts in the realm of cybersecurity have been undertaken to safeguard servers against malicious attackers and hackers. However, conventional cybersecurity methods have shown shortcomings in protecting servers from various forms of unauthorized external traffic. Addressing this, the development of an Intrusion Detection System (IDS) tailored to the Internet of Things (IoT) architecture becomes crucial. Rigorous literature reviews were conducted to delve into different machine learning techniques, neural network models, and optimization algorithms, aimed at pinpointing gaps, challenges, and subsequently formulating accurate and effective machine learning algorithms for precise intrusion detection.

Specifically, in the artificial neural network (ANN) model, two strategies were formulated for detecting Distributed Denial of Service (DDoS) attacks: the Back Propagation Neural (BPN) and the Multi-layer Perceptron (MLP) methods. To enhance their performance, a novel hybrid optimization algorithm was introduced. This algorithm combines the exploitation capabilities of the Harris Hawks Optimization (HHO) technique with the exploration capabilities of Particle Swarm Optimization (PSO). By doing so, common limitations observed in traditional algorithms—such as local stagnation issues, delayed convergence concerns, and the challenges of getting trapped in local or global optima—are effectively managed by this innovative hybrid optimization approach.

Furthermore, the proposed hybrid HHO-PSO algorithm goes beyond feature selection. It's harnessed to fine-tune the neural network models by adjusting the weight and bias coefficients. This comprehensive approach demonstrates a concerted effort to bolster intrusion detection accuracy and efficiency in cloud computing environments, particularly within the context of IoT architecture

Keywords— Cloud Computing, Neural Network (CNN) , Hadoop.

1. INTRODUCTION

The rapid expansion of the information technology sector has led organizations of all sizes and types to offer online services directly to consumers. The advent of cloud computing and the Internet of Things (IoT) has revolutionized service provision, enabling on-demand access for users via the internet [1]. This paradigm grants access to extensive data storage capabilities accessible globally at any moment. Both public and private sectors are seamlessly integrated through a shared service infrastructure. However, this diversity also creates opportunities for illicit cyber-attacks, including Denial of Service (DoS), Distributed Denial of Service (DDoS), Drive-by attacks, SQL injection attacks, Man-in-the-Middle (MitM) attacks, and more. Among these, DDoS attacks represent a particularly disruptive intrusion technique, where a network is overwhelmed by a sudden flood of attack packets, often using compromised devices to induce traffic congestion. Differentiating between legitimate and malicious traffic in such situations poses a significant challenge. Therefore, the focal point of our proposed research is the development of an intelligent intrusion detection system specifically designed to identify DDoS attacks in the cloud computing environment.

Cloud computing is characterized by the provisioning of software and hardware computing resources over the internet, tailored to user demand and payment based on usage. This diverse landscape unites various companies and organizations within a shared resource ecosystem. Employing dynamic scaling strategies, cloud computing ensures high reliability and adaptability in service delivery. Instead of traditional client-server architecture, it leverages virtualization, enabling seamless transition of servers into distinct virtual machines [2]. The key advantage of cloud computing lies in the communal sharing of resources, which obviates the need for users to purchase third-party applications, facilitating resource access at any given time.

2. MODEL TO DDOS ATTACK DETECTION

we have developed neural network models utilizing straightforward feed-forward neural network architectures—specifically, the Back Propagation Neural Network and the Multilayer Perceptron network. To enhance the performance of these models, we have employed a wrapper-based strategy.

Furthermore, we have introduced a novel swarm intelligence-based hybrid optimization algorithm, amalgamating the strengths of Particle Swarm Optimization (PSO) and Harris Hawks Optimization (HHO). Traditional population-based optimization algorithms often grapple with suboptimal performance across diverse problem statements, plagued by issues such as local optimal entrapment, delayed convergence, global optimal strapping, and premature convergence. By combining these two distinct algorithms, we effectively address these limitations, ensuring that the models encounter no such issues. The PSO algorithm excels in exploration capabilities, while HHO demonstrates superior exploitation capability. This synergy empowers the hybrid algorithm to adeptly tackle the aforementioned challenges.

Furthermore, the proposed hybrid algorithm serves a critical role in tuning the weight and bias coefficients of the neural network models. Typically, these models initiate with random weight and bias vectors that evolve during training. However, this random initialization can lead to delayed convergence and subpar performance. By initializing the models with optimal weight and bias vectors, performance can be significantly elevated. The proposed HHO-PSO algorithm is harnessed to selectively determine these initial weight and bias vectors for the neural network models, optimizing their effectiveness and overall performance.

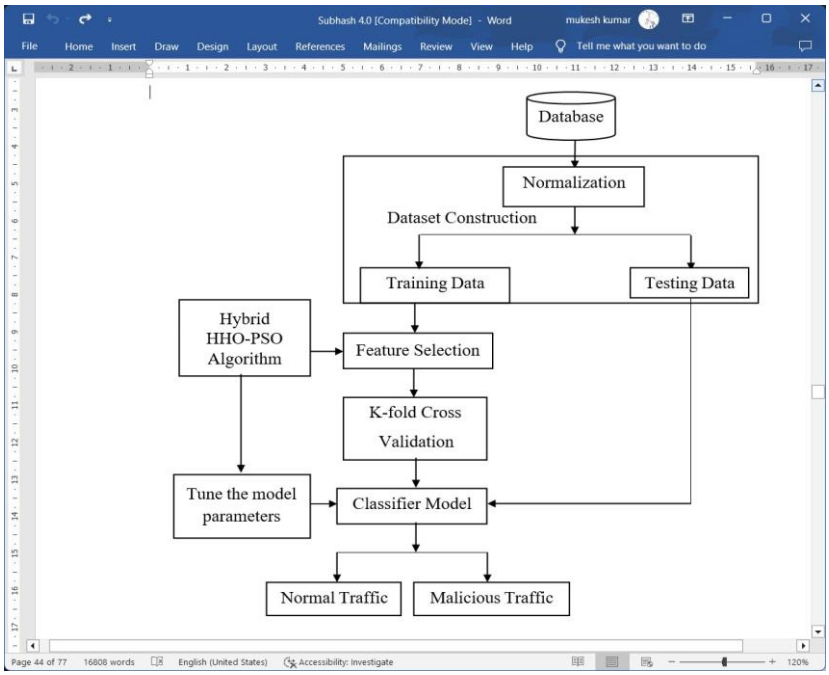

Fig 1: Machine Learning based Distributed Denial of Service Attack Detection Model

MapReduce programming offers a potent solution for managing extensive calculations by leveraging flexibility and reliability features. This approach capitalizes on both map and reduce stages to achieve parallel execution of tasks. The mapper task processes individual input splits and generates key-value pairs. These key-value pairs differ from the input and are provided to the mapper function. The resulting key-value pairs from each mapping task are organized based on their keys. Subsequently, the reduce task manages and processes each key, merging all associated values linked to that particular key.

3. RESULT AND DISCUSSION

The proposed data mining algorithms have been effectively trained using the training dataset within a 10-fold cross validation setup. Throughout the training process, the essential features for C4.5, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) are identified individually for each fold. The results of this feature selection process are presented in Table 4.2 to Table 4.5. The selected features' frequencies are also documented.

The introduced approach demonstrates proficient performance with sizable datasets, resulting in improved execution times when compared to the conventional K-means clustering technique. The KM-HMR method builds upon the MapReduce programming model, incorporating both mappers and reducers. In this scheme, each mapper task is responsible for assigning the items or data points closest to a cluster during the iteration. Subsequently, the reducer task processes the files generated by the mappers, traversing through the list of clusters and associating each object to its respective cluster [16, 17].

3.1 COMPARISON BETWEEN SIMPLE K-MEAN AND PROPOSED KM-HMR

The proposed Data mining algorithms are successfully trained with the training dataset of 10-fold cross validation. During the process of training, the essential features of C4.5, KNN and SVM are identified for each fold and presented. The frequency of selected features of C4.5, KNN and SVM during the 10-fold cross validation and their corresponding accuracy are shown in Figure 4.2.

Table 4.2 Feature Selection by the proposed C4.5 IDS model

Fold	Selected Features	Fold	Selected Features
#1	F3,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36,F37,F38	#6	F4,F5,F8,F10,F12,F23,F25, F29,F30,F36,F37,F41
#2	F4,F5,F6,F8,F10,F11,F12,F23, F25, F29,F30,F35,F36,F37	#7	F4,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36
#3	F3,F4,F5,F8,F10,F12,F23,F25, F29, F30,F35,F36,F37	#8	F4,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36,F37,F38
#4	F3,F4,F5,F7,F8,F10,F12,F23, F25, F29,F30,F35,F36	#9	F4,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36,F37
#5	F4,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36,F37	#10	F4,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36,F37

Table 4.3 Feature Selection by the proposed KNN IDS model

Fold	Selected Features	Fold	Selected Features
#1	F4,F5,F8,F10,F12,F23,F29,F30, F33,F35,F36,F37,F39	#6	F4,F5,F6,F8,F10,F12,F23, F25,F26,F29,F30,F33,F35, F36,F37,F39
#2	F4,F5,F8,F10,F12,F23,F25,F26, F29,F30,F33,F35,F36,F37,F39	#7	F4,F5,F8,F10,F12,F23,F25, F26, F29,F30,F35,F36,F37, F38,F39
#3	F4,F5,F6,F8,F10,F12,F25,F29, F30,F33,F35,F36,F37,F38,F39	#8	F4,F5,F8,F10,F12,F23,F25, F26, F29,F30,F33,F35,F36, F37,F39
#4	F4,F5,F6,F8,F10,F12,F23,F25, F26, F29,F30,F33,F35,F36,F39	#9	F4,F5,F8,F10,F12,F23,F25, F26,F29,F30,F35,F36,F37, F39
#5	F4,F5,F6,F8,F10,F12,F23,F25, F26,F29,F30,F33,F35,F36,F37, F38,F39	#10	F4,F5,F8,F10,F12,F23,F25, F26,F29,F30,F35,F36,F37, F39

The better accuracy has been achieved while feeding the models with the feature subset that has frequency of occurrence between 8 and 9**.** So, the feature that has the frequency of occurrence between 8 and 9 is considered to train the model and the corresponding performance is analyzed.

Table 4.4 Feature Selection by the proposed SVM IDS model

Fold	Selected Features	Fold	Selected Features
#1	F4,F5,F8,F10,F12,F23,F25,F26 ,F29,F30,F35,F36,F37,F38	#6	F4,F5,F8,F10,F12,F23,F29, F30,F35,F36,F37
#2	F4,F5,F8,F10,F12,F23,F26, F29,F30,F35,F36,F39	#7	F4,F5,F8,F10,F12,F23,F26, F29,F30,F35,F37, F38
#3	F4,F5,F8,F10,F12,F23,F26, F29,F30,F34,F35,F36,F38	#8	F4,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36,F37,F39
#4	F4,F5,F8,F10,F12,F23,F26, F29,F30,F35,F36,F38	#9	F4,F5,F8,F10,F12,F23,F26, F29,F30,F35,F36
#5	F4,F5,F8,F10,F12,F23,F26, F29,F30,F35,F36,F38	#10	F4,F5,F8,F10,F12,F23,F25, F29,F30,F35,F36

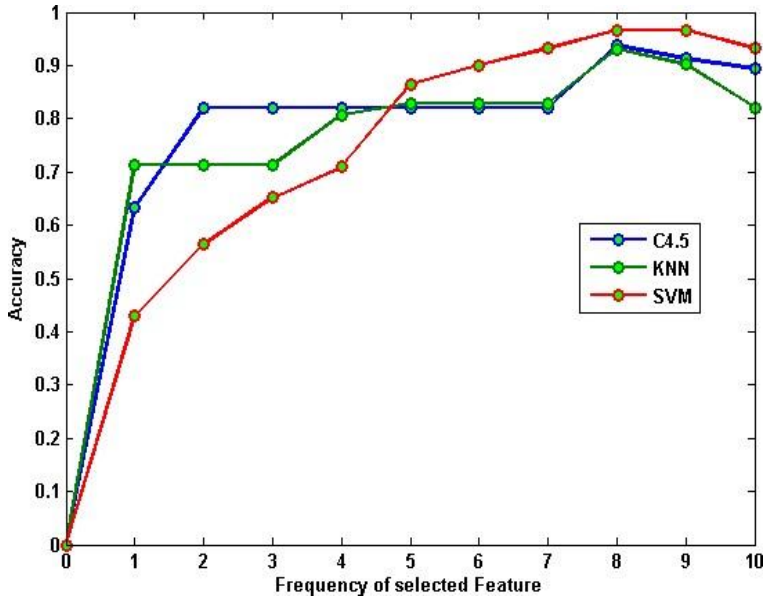

Figure 4.2 Accuracy Vs selected features of C4.5, KNN and SVM

4. CONCLUSION

The validity of the proposed model is confirmed through an independent dataset, employing specified performance metrics, and then compared against the performance of existing models. By integrating these two algorithms, the benefits of both are harnessed, leading to superior outcomes compared to traditional methods. As a result, this study combines the C4.5 classifier algorithm with SVM and KNN models, and the results conclusively demonstrate that the proposed SVM-based classifier model outperforms all other models in terms of intrusion detection performance. Additionally, it achieves this with a minimal number of feature subsets compared to other models.

While the proposed classifier models yield improved classification outcomes, it's noteworthy that the SVM classifier algorithm can face challenges when dealing with large datasets. This limitation can be effectively addressed by introducing the Artificial Neural Network (ANN) algorithms. The unique nature of ANN lies in its ability to handle extensive datasets, attributed to its proficient learning mechanism that emulates the human brain's learning process.

5. REFERENCES

[1] Aamir, M & Zaidi, SMA 2019, 'DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation', International Journal of Information Security,vol.18, no.6, pp.761-785.

[2] Abawajy, J, Huda, S, Sharmeen, S, Hassan, MM & Almogren, A 2018, 'Identifying cyber threats to mobile-IoT applications in edge computing paradigm', Future Generation Computer Systems, vol.89, pp.525-538.

[3] Abeshu, A & Chilamkurti, N 2018, 'Deep learning: The frontier for distributed attack detection in fog-tothings computing', IEEE Communications Magazine, vol. 56, no. 2, pp.169-175.

[4] Abubakar, R, Aldegheishem, A, Majeed, MF, Mehmood, A, Maryam, H, Alrajeh, N, Carsten, M & Jawad, M 2020, 'An Effective Mechanism to Mitigate Real-time DDoS Attack Using Dataset', IEEE Access.

[5] Adhikary, K, Bhushan, S, Kumar, S & Dutta, K 2020, 'Hybrid Algorithm to Detect DDoS Attacks in VANETs', Wireless Personal Communications, pp.1-22.

[6] Agarwal, S, Tyagi, A & Usha, G 2020, 'A Deep Neural Network Strategy to Distinguish and Avoid Cyber-Attacks', In Artificial Intelligence and Evolutionary Computations in Engineering Systems, Springer, pp. 673-681.

[7] Aggarwal, P & Sharma, SK 2015, 'Analysis of KDD dataset attributes- class wise for intrusion detection', Procedia Computer Science, vol. 57, pp.842-851.

[8] Ahanger, TA 2017, 'An effective approach of detecting DDoS using Artificial Neural Networks', In 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), IEEE, pp. 707-711.

[9] Ahmed, AA, Jabbar, WA, Sadiq, AS & Patel, H 2020, 'Deep learning- based classification model for botnet attack detection' 'Journal of Ambient Intelligence and Humanized Computing', pp. 1-10.

[10] Ali, O & Cotae P 2018, 'Towards DoS/DDoS Attack Detection Using Artificial Neural Networks', In 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), IEEE, pp. 229-234.

[11] Ali, S & Li, Y 2019, 'Learning Multilevel Auto-Encoders for DDoS Attack Detection in Smart Grid Network', IEEE Access, vol. 7, pp.108647-108659.

[12] Ali, SHA, Ozawa, S, Ban, T, Nakazato, J & Shimamura, J 2016, ' A neural network model for detecting DDoS attacks using darknet traffic features', In 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, pp.2979-2985.

[13] Ali, U, Dewangan, KK & Dewangan, DK 2018, 'Distributed Denial of Service Attack Detection Using Ant Bee Colony and Artificial Neural Network in Cloud Computing', In Nature Inspired Computing, Springer, pp. 165-175.

[14] Aljumah, A & Ahamad, T 2016, 'A novel approach for detecting DDoS using artificial neural networks', International Journal of Computer Science and Network Security, vol. 16, no.12, pp.132-138.

[15] Aljumah, A, 2017, 'Detection of distributed denial of service attacks using artificial neural networks, International Journal of Advanced Computer Science and Applications, vol. 8, no. 8.

[16] Alkasassbeh, M, Al-Naymat, G, Hassanat, A & Almseidin, M 2016, 'Detecting distributed denial of service attacks using data mining techniques', International Journal of Advanced Computer Science and Applications, vol. 7, no.1, pp.436-445.

[17] Alzahrani, S & Hong, L 2018, 'Detection of distributed denial of service (DDoS) attacks using artificial intelligence on cloud', In 2018 IEEE World Congress on Services (SERVICES). IEEE, pp. 35-36.

[18] Asad, M, Asim, M, Javed, T, Beg, MO, Mujtaba, H & Abbas, S 2019, 'DeepDetect: Detection of distributed denial of service attacks using deep learning', The Computer Journal.

[19] Baek, UJ, Ji, SH, Park, JT, Lee, MS, Park, JS & Kim, MS 2019, 'DDoS Attack Detection on Bitcoin Ecosystem using Deep-Learning', In 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), IEEE, pp. 1-4.

[20] Barış K, Çağatay Y, Taha YC, Bülent S &Ali TC 2018, 'A Bayesian change point model for detecting SIP-based DDoS attacks', Digital Signal Processing. vol. 77, pp. 48–62.

[21] Bawany, NZ, Shamsi, JA & Salah, K 2017, 'DDoS attack detection and mitigation using SDN: methods, practices, and solutions', 'Arabian Journal for Science and Engineering' vol. 42, no. 2, pp.425-441.

[22] Behal, S, Krishan, K &Sachdeva, M 2018, 'D-FACE: An anomaly based distributed approach for early detection of DDoS attacks and flash events', Journal of Network and Computer Applications, vol. 111, pp.49–63.