International Journal of Communication Networks and Information Security
2024, 16(S1)
ISSN: 2073-607X, 2076-0930
https://https://ijcnis.org/ **Research Article**

Cloud Based DDOS Attack Detection in a Distributed and Collaborative Manner Using Deep Neural Networks

Aditya Kumar Shukla[1], Ashish Sharma[2], Sandeep Singh Sengar[3]

[1,2]Department of Computer Engineering and Applications, GLA University, NH#2, Delhi Mathura Highway, Post Ajhai, Mathura (UP) India.
[3]Department of Computer Science; Cardiff School of Technologies, Cardiff Metropolitan University, UK.

Email ID: [1]uraditya@gmail.com , [2]ashishs.sharma@gla.ac.in , [3]sssengar@cardiffmet.ac.uk

ARTICLE INFO

Received: 19 Apr 2024
Accepted: 28 Aug 2024

ABSTRACT

Attacks known as DDOS (Distributed Denial of Service) are now a significant threat to the Internet's availability. Traditional methods of detecting DDOS attacks are limited in their effectiveness due to the heterogeneous nature of cloud data and application deployment. In today's time, applications are deployed in different containers and nodes over the cloud, and application solutions are deployed in different regions even if they are not fixed with a single cloud service provider. The biggest challenge is data privacy preservation while fighting DDOS attacks while applications are deployed in a distributed manner over the cloud. This study suggests a combination of deep neural networks and federated learning strategies in order to identify distributed denial of service (DDOS) attacks in an environment of heterogeneous cloud service providers. This strategy leverages locally trained models to detect anomalies across different cloud nodes. The models are trained using features extracted from different cloud service providers' logs and traffic data. This strategy is aimed at providing more secure and robust detection of DDOS attacks compared to traditional methods and preserving cloud data privacy. Utilizing several criteria, including accuracy, exactness, recall, and F1-score, the combination of CNN(convolutional neural network) and the federated learning-based detection model is assessed. The outcome of the experiment shows the efficacy of the suggested method in detecting DDOS attacks with high accuracy. This strategy can be used to detect DDOS attacks in an environment of heterogeneous service providers, providing a more secure and robust detection framework.

Keywords: DDOS Attack Detection, Heterogeneous Cloud Service Providers, Federated Learning Strategy, Cloud computing, Deep Neural Networks etc.

INTRODUCTION

Nowadays, society and the business sectors significantly rely on the innovative computation technique known as cloud computing. it allows users to utilize computation resources remotely on any computer with an internet connection [1]. Cloud computing is come to realty with the help of well-connected and secured communication channels, whose role is to offload-and download the data during computation, which make the system to more exposed and flexible to both users and intruders [2]. Even though cloud service providers built a cyber-attack malware at every possible node of the system, intruders are come up with new techniques to attack and stole the data. Which in turn raises the need of a well-trained Global intuition detection system to prevent from such unwanted events [3].

Cloud servers get compromised, become online zombies, and cause well-known bots and crises. Due to their secrecy, abundance, and affordable cost, hackers like to use these zombies to assault other entities. DDoS (distributed denial of service) attacks have thus grown to be a severe security risk on the World Wide Web.

Attacks that cause a distributed denial of service (DDoS) are a form of cyberattack that uses multiple computers to target a single victim, usually with the intent of overwhelming the victim's network with traffic and thereby disrupting its service. DDoS attacks can be used for malicious purposes such as vandalism, extortion, or even political protest [10]. In cloud environments, where resources are shared among multiple users, Data security and service availability are seriously threatened by DDoS assaults. [6]. A Heterogeneous System is one that consists of multiple types of systems, such as a variety of hardware, software, or other components. A heterogeneous system can be thought of as a combination of two or more different types of systems. A cloud environment is a type of heterogeneous system that combines multiple computers and other resources to provide services to users [11].

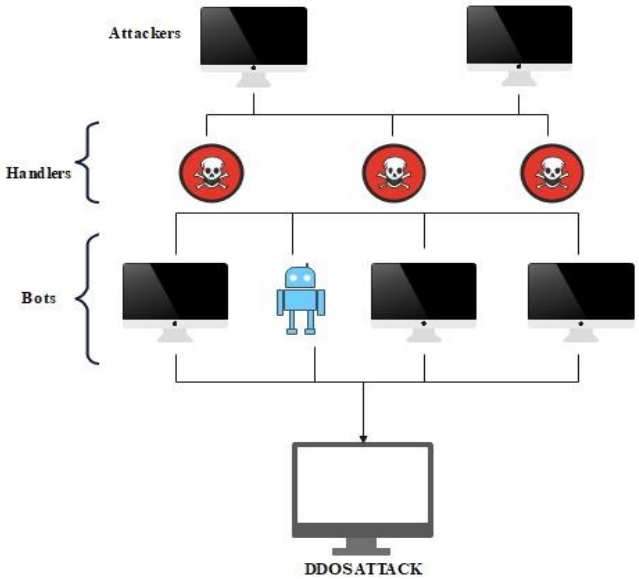


Figure 1 DDOS Attack

On the Internet, an intruder organizes bots, whereas a defender engages in solitary DDoS defence[Figure 2]. Attackers purchase inexpensive nodes of networks of bots to assault a target at will, making spectacular gains from solitary defenders due to their resource limitations and massive size.

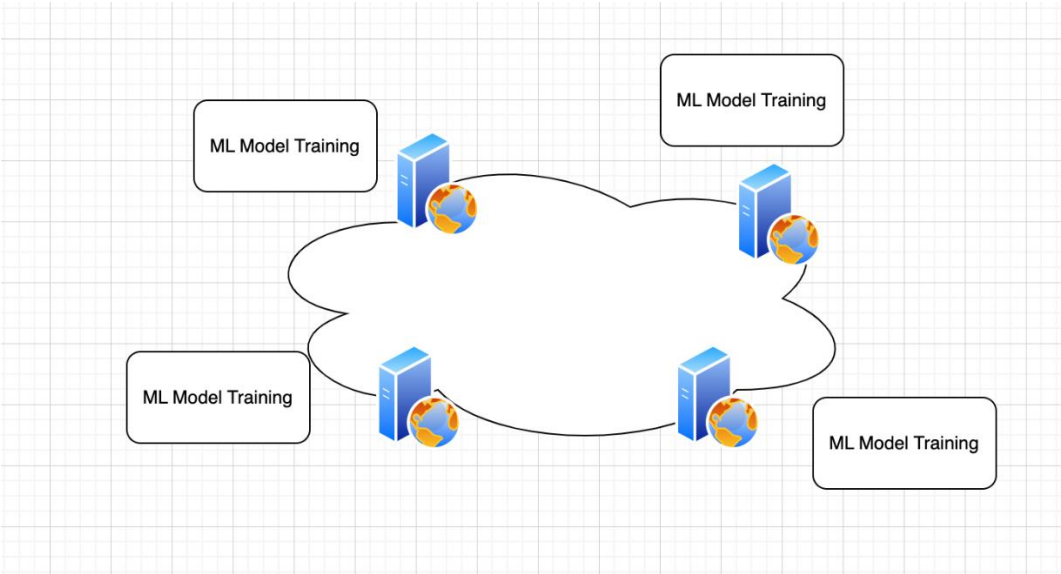


Figure 2 Defender fights against DDoS individually

DDOS attacks can be difficult to mitigate in cloud environments because of the sheer scale of the attack [4]. Since a cloud environment is composed of many different components, it can be difficult to identify and stop the attack traffic on each component. Moreover, because the cloud is a shared resource, it might be challenging to prevent attack traffic from spreading from one component to another [5].

The current DDoS detection techniques require global evaluation of traffic information from many sources, but it is evident that such collection of information is not beneficial to source users' protection of privacy [6,7]. Traditional machine learning methods call for a model to learn from a large number of training samples, which can occasionally be very challenging to gather due to privacy concerns[Figure 3].

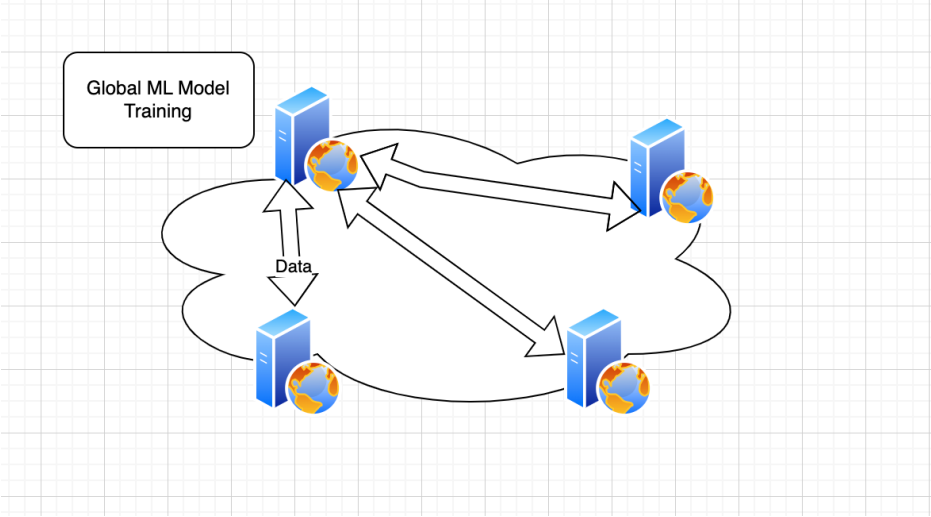


Figure 3 Global learning of traffic information from many sources

Federated learning can be used to detect DDOS attacks in a heterogeneous service provider environment. With just the model parameters being sent to a secure server, each service provider trains its own local model using only its own set of data. The server can aggregate the model's inputs from all the service providers and use them to create a global model, which can

be used to detect anomalies that indicate a DDOS attack [8,9]. This approach is more secure, as no sensitive data needs to be transferred and shared between the service providers.

FL is a method of iteration where each iteration allows for the improvement of the entire ML/DL model. The FL server chooses a group of clients who will take part in the learning procedure and distributes the present global model to the group at the start of every cycle.

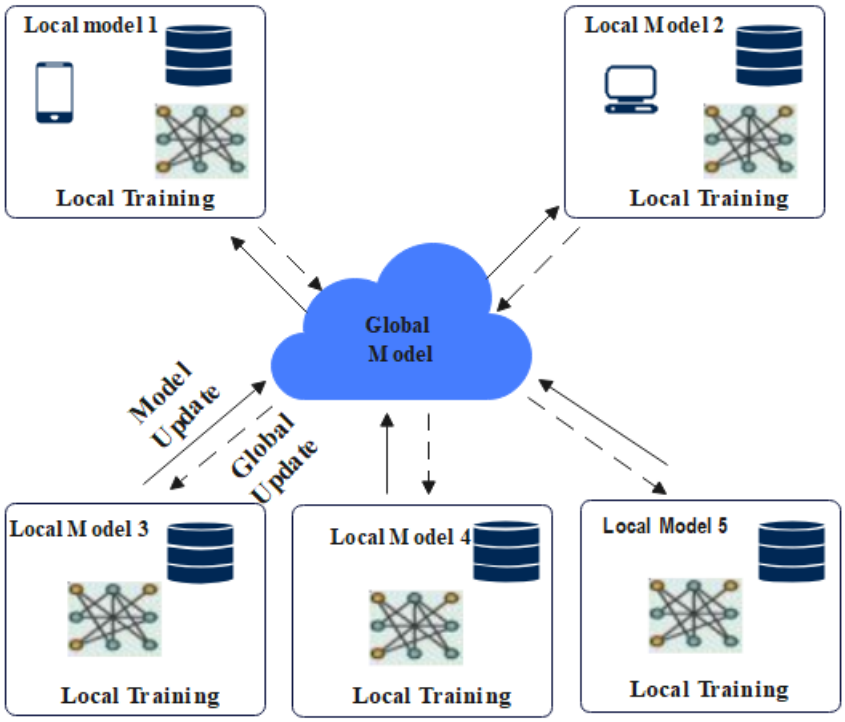


Figure 4 General Federated Learning Architecture

Every client utilizes their own information to perform regional learning after getting a global model. These clients subsequently send the parameters they have learned from the fresh model back to the FL servers for global aggregating. Up until the level of performance you want is attained, the procedure is done numerous times. In conclusion, there are two key stages in a collaborative learning instance: local updating and global aggregation. This highlights how FL enables users to gain access to the records of other users while submitting their own privacy-sensitive information to a centralized server.

An input layer, hidden layers, and an output layer make up any deep neural network. Convolutional neural networks (CNN), a type of deep neural network, are used in the classification of images, which is significant. The novelty in this study is that it combines FL and CNN for DDoS detection in cloud computing, a technique that has not been explored in previous research.

Organization of the Study

The work may be broken down into the following sections: Literature reviews are provided in Section II, an explanation of the methodology that underpins the recommended algorithm is provided In Section III, Section IV presents the study's results, and Section V offers a summary and a conclusion. Chapter 5 is where the novel is finished.

LITERATURE REVIEW

Petrakopoulos, Vasilis [8], describes the problems to build global model in the heterogeneous system. Even though cloud computing is emerging market now days, there are only few players in the market globally, such as Amazon Web Services, Google Cloud Platform, Microsoft Azure, etc. as they are competitors they would not be interested to share their data among themselves.

A. Cheema, M. Tariq [9] had presented the comprehensive classification of DDOS attacks, as well as identified the reasons of attackers behind these attacks, as well as the repercussions of these attacks.

R. Doriguzzi-Corin and D. Siracusa [10] looked at how the procedure of the FL converges in the ever-changing field of cybersecurity situations, where the learned model is located must be brought up to date regularly with all new threat characteristics to be providing every one of federation participants in the most recent recognition tools.

Tang, Zhongyun [11] described In order to broaden the scope of a cross-silo setup for a network intrusion detection system (NIDS) based on flow, the stacked-unsupervised federated learning (FL) technique has been proposed. **J. Kim** [12] had developed DDOS attacks are the primary emphasis of this DL-based intrusion model. In this paper, author evaluated several intrusion detection systems using the most widely-used dataset available, the KDD CUP 1999 dataset (KDD) (IDS).

A. A. E. Cil [13] had suggested DDOS attacks may be detected by analyzing a subset of packets captured from network traffic using a deep neural network (DNN) model. There is a process called feature extraction as well as categorization of the structure that is carried out before the dataset can be used to train the model.

X. Sáez-de-Cámara [14] and **S. Arisdakessian** [15] said DDOS attacks can be especially dangerous, due to the cloud is composed of many different components, and DDOS attacks can be used to target one or more of these components in order to disrupt the entire system. For example, an attacker could target a single server in the cloud, causing it to become overwhelmed with traffic and thereby disrupting the entire cloud system. Furthermore, because the cloud is a shared resource, an attack on one component can have a ripple effect, affecting other components as well

V. Mothukuri [16] DDOS attacks can also be difficult to detect in cloud environments because the attack traffic is often distributed across multiple components of the cloud, making it difficult to identify the source of the attack. **G. De Carvalho Bertoli** [17] put it another way, because to its properties, ML/DL may be taught utilizing distributed data over a number of iterations and multiple servers and devices.

N. N. Dao et al [18] explained that The server FL chooses a set of customers to take part in the procedure for learning at the beginning of each cycle and gives them its most recent global model. Each client utilizes its own data for local training after getting the global model.

Li, Kun, et al. [19] discussed that Two crucial phases—local updating and global aggregation—are present in a federated learning scenario, to sum up. This exemplifies how FL enables users to see the data of other users without revealing their own sensitive personal information to a central server.

Dora, V. Raghava Swamy, and V. Naga Lakshmi [20] developed a DDoS detection model using deep learning by combining a convolutional neural network (CNN) and an optimized long short-term memory (LSTM), known as CNN-O-LSTM. Lv, Dingyang, et al. [21] offer a FLDDoS system that combines federated learning (FL) with neural networks to counter DDoS attacks.

The studies that have been mentioned so far all used two methods. 1) Conventional ML, where a single computer or central server handles data storage and model training. As a result, prior to training, all the data must be gathered in one location. 2) Employed edge or fog computing with distributed machine learning.

Contribution of This Study:

1. This work demonstrates that without sacrificing accuracy, the federated learning approach can accomplish privacy-protected DDoS detection in the cloud.
2. Apply CNN machine learning methods for DDoS assault categorization; only weights will be distributed among the global server and local client; models are trained using the shared global weight in a distributed way.
3. Compared to other ML models and the centralised ML training method, the accuracy is more accurate.
4. The suggested model's contribution is to demonstrate that using deep neural networks in a federated context does not compromise performance. Therefore, using the FL technique, we can create reliable classifiers in the cloud even if there are privacy concerns with data collecting.

METHODOLOGY

The security of cloud computing has become a crucial concern due to its rise in demand. Numerous threats that limit the use of cloud computing have been found. One of the biggest obstacles to cloud computing is spotting DDOS attacks. The proposed model's contribution is to demonstrate that using deep neural networks (convolutional neural network) in a federated context does not compromise performance. Therefore, using the FL technique, we can create reliable classifiers in the cloud even if there are privacy concerns with data collecting.

Combination of Federated Learning and Convolutional Neural Network

An approach to FL known as FFL (also known as features-based FL) is one in which data from many domains are combined in order to train a single global model. In this case, various client datasets may include the same observations but have distinct characteristics. FL is a distributed training framework that protects privacy and consists of several participants jointly training a single ML model on their local datasets. The initialization of the global model by a centralized entity kicks off the iterative training process. In every round of dialogue, Among N participants, a subset is given the most current global model. The model is then trained by each participant k by executing many rounds of using micro batches of its local dataset D_k and stochastic gradient descent (SGD). The local instruction produces a number of weight-updating vectors, that are forwarded to the server. And the next stage is the server's model aggregation, which is commonly accomplished via aggregated by weight as described in Eq. Following that, the process is performed again till convergence of the model.

$$\theta_{t+1} = \theta_t + \sum_{K=1}^{N} \frac{|D_K|}{|D|} \Delta\theta_K^{t+1}$$

The local training generates a number of weights for the neural network, which are then sent to the server to be integrated. This is often done via weighted aggregation, which is a technique that is rather popular. After then, the process is repeated as many times as necessary until the model reaches a stable state. The weights of the neural network were updated from the local models to the global model using the federated averaging method. The Federated Averaging

technique is created to address this issue, which is one of the key challenges involved in making this practicable. The problem involves lowering the pace at which information is sent between the global model and the regional models.

Algorithm 1 Deep neural network based on federated learning, with Targeting updates from K Models per Round
The global model is put into action as follows:
Set up: ω_o
Do this **for each round**
// Generate updates for 1.3K eligible local models.
// Request updates from K local models (indexed 1.K).
$\left(\Delta^k, n^k\right) =$ Local model update (ω) from local model $k \in [K]$
$\overline{\omega_t} = \sum_k \Delta^k$ // sum of the weighed updates
$\overline{n_t} = \sum_k n^k$ //sum of weights
$\Delta_t = \Delta_t^k / \overline{n_t}$ // Average update
$\omega_{t+1} \leftarrow \omega_t + \Delta_t$
Local model update (ω):
$\mathrm{B} \leftarrow$ (separated into tiny batches of local data)
$n \leftarrow \lvert \mathrm{B} \rvert$ //Update weight
$\omega_{init} \leftarrow \omega$
$k \in [K]$ for batch $b \in \mathrm{B}$ do
$\omega \leftarrow \omega - \eta \nabla \ell(\omega; b)$
$\Delta \leftarrow n \cdot (\omega - \omega_{init})$ // weighed update
//Note Δ is susceptible to compression greater than ω return (Δ, n) to global model

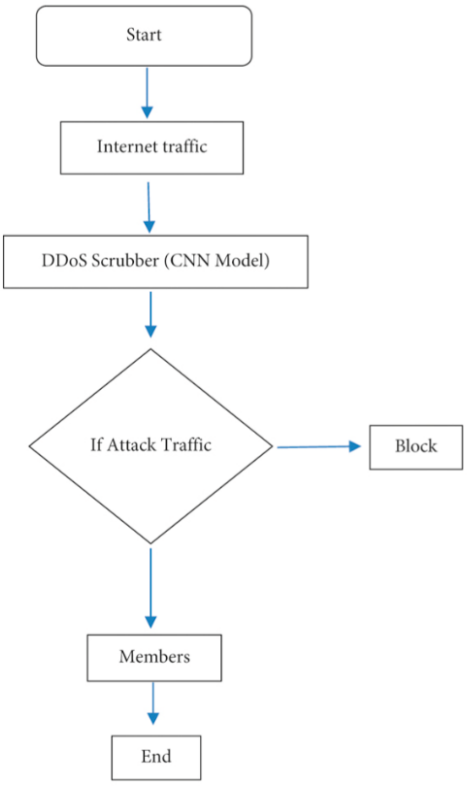


Figure 5 CNN based DDOS classification

Convolutional neural networks (CNN) use an assembly of layers that are convolutional and pooling in the initial stages, a number of fully connected layers in the last stage, and then a classifier based on softmax to categorize the data that is input. Distributing parameters and having a few connections are two of CNN's key characteristics. The characteristics employed in one section can also be convolved over the entire system in sharing parameters after the extraction of features.

This study made use of the NSL-KDD standard DDoS data set. The following table shows the distribution of the data.

Table 1: Distribution of DDoS data in NSL-KDD

	Initial records	**Identical records**	**Rate of Reduction(%)**
Assaults	3925650	262178	93.32
Ordinary	972781	812814	16.44
Sum	4898431	1074992	78

RESULTS

In this research study, we proposed a DDoS detection system based on the combination of CNN and federated learning for the precise attack identification and categorization within heterogeneous networks. To increase classification accuracy rates, FL implementation is a component of our suggested technique computational shares resources using training for on-device. The results of our comparison testing show that our proposed technique performs better than non-FL versions of intrusion detection methods.

This section displays the findings of the study of our suggested method (FL) in comparison to non-FL implementation. The capacity of several computing instances is shared, and parallel computing minimizes the amount of training time needed to reach peak performance.

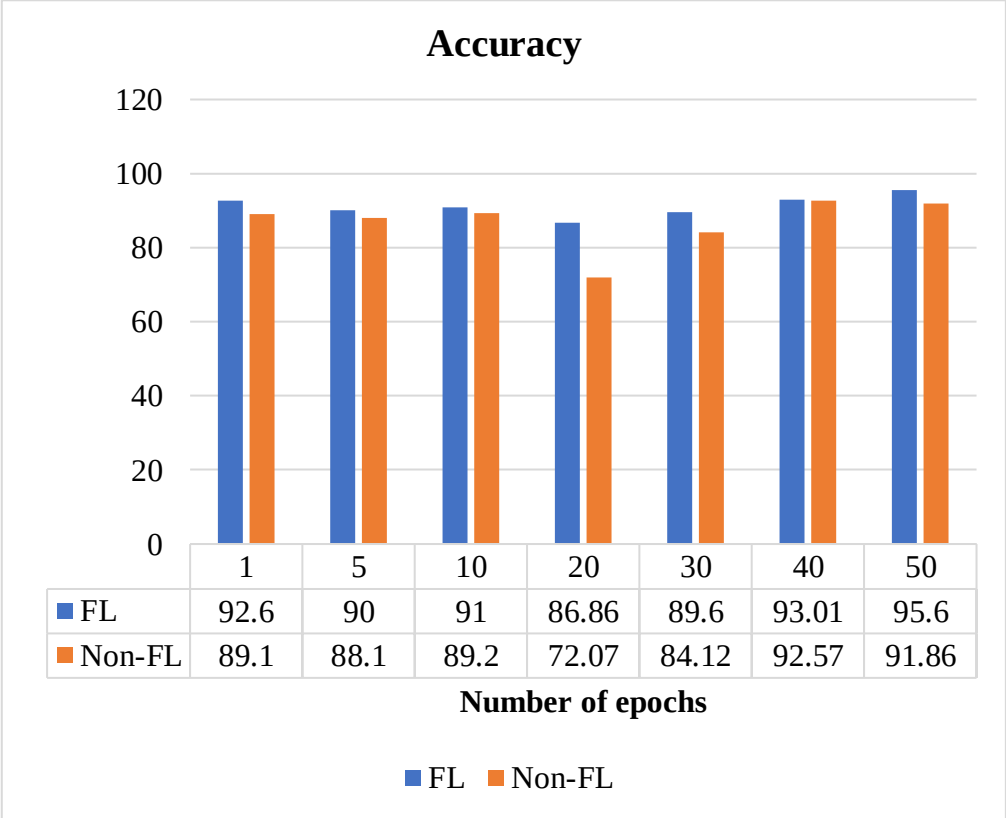


Figure 6 Accuracy comparison between Distributed and Centralized Approach

Using the dataset and assessment criteria indicated in the above image, we compare the accuracy of our suggested technique with a non-FL variant. For epoch 50, the accuracy values for FL and non-FL are 95.6 and 91.86, respectively. When compared to the non-FL variant, our recommended FL version succeeds better precision with fewer epochs.

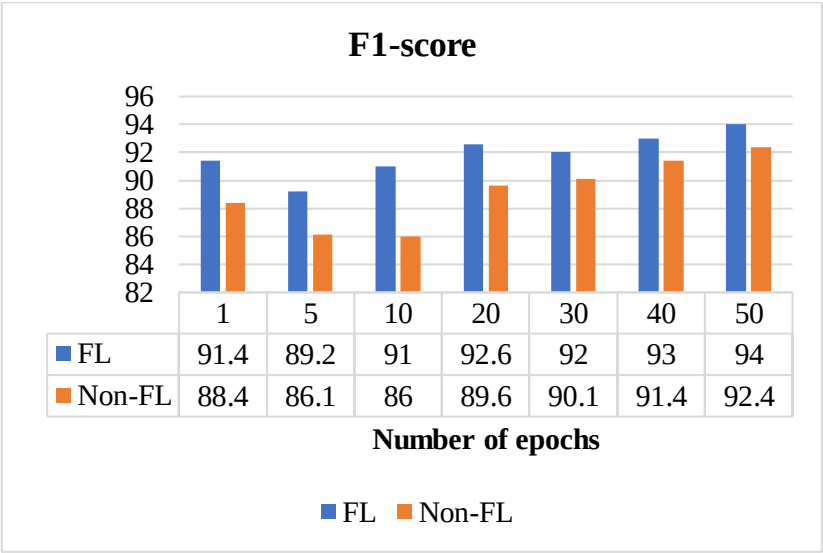


Figure 7 F1 Score comparison between Distributed and Centralized Approach

The F1-score of both FL and Non-FL with different epochs are presented in the above figure. It can be seen that epoch 50 has higher F1-score value than epoch1 for both FL and Non-FL versions. The F1-score value of epoch 50 for FL is 94 and non-FL is 92.4.

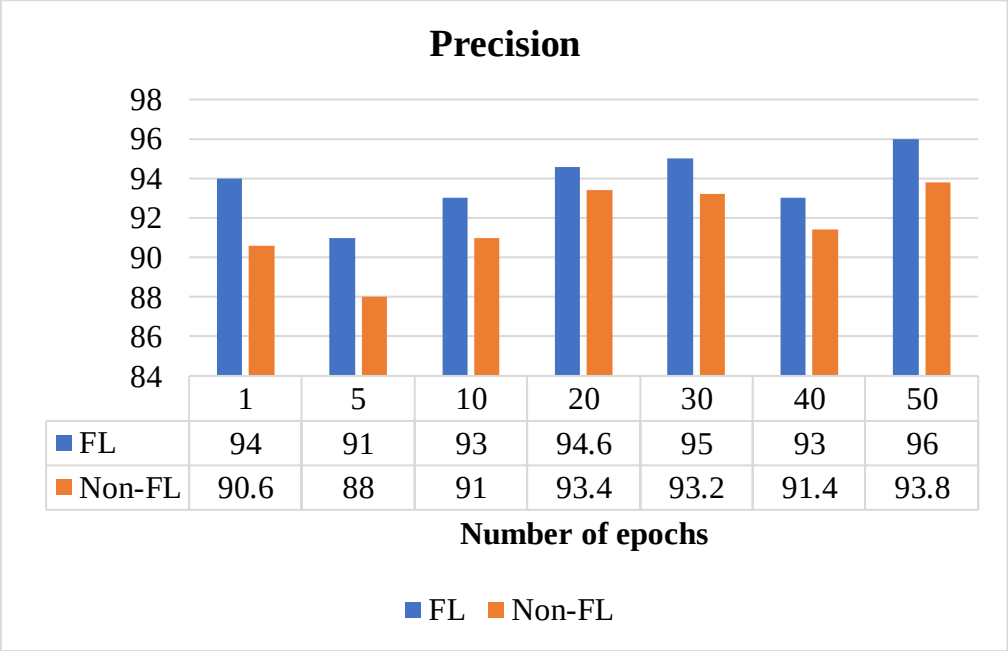


Figure 8 Precision comparison between Distributed and Centralized Approach

In order to evaluate the precision of the approach that we have suggested, We contrast it with a non-FL variant that has a number of epochs that are different from those shown in the above figure. In epoch1, the precision value for FL is 94 and non-FL is 90.6, which is rapidly increased to higher precision value for both FL and Non-FL versions in epoch50.

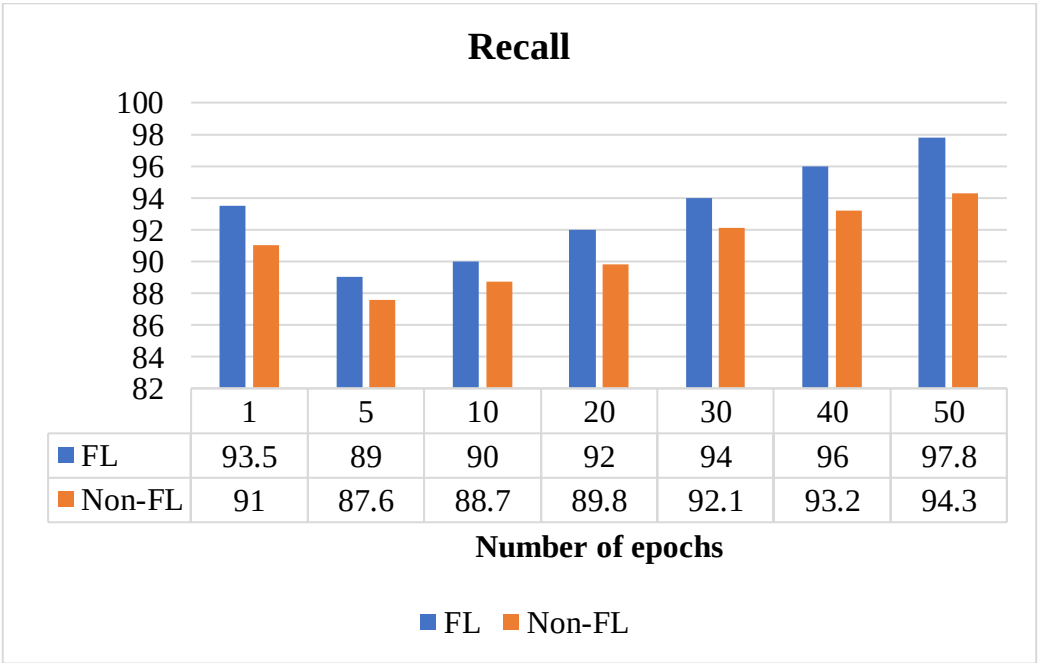
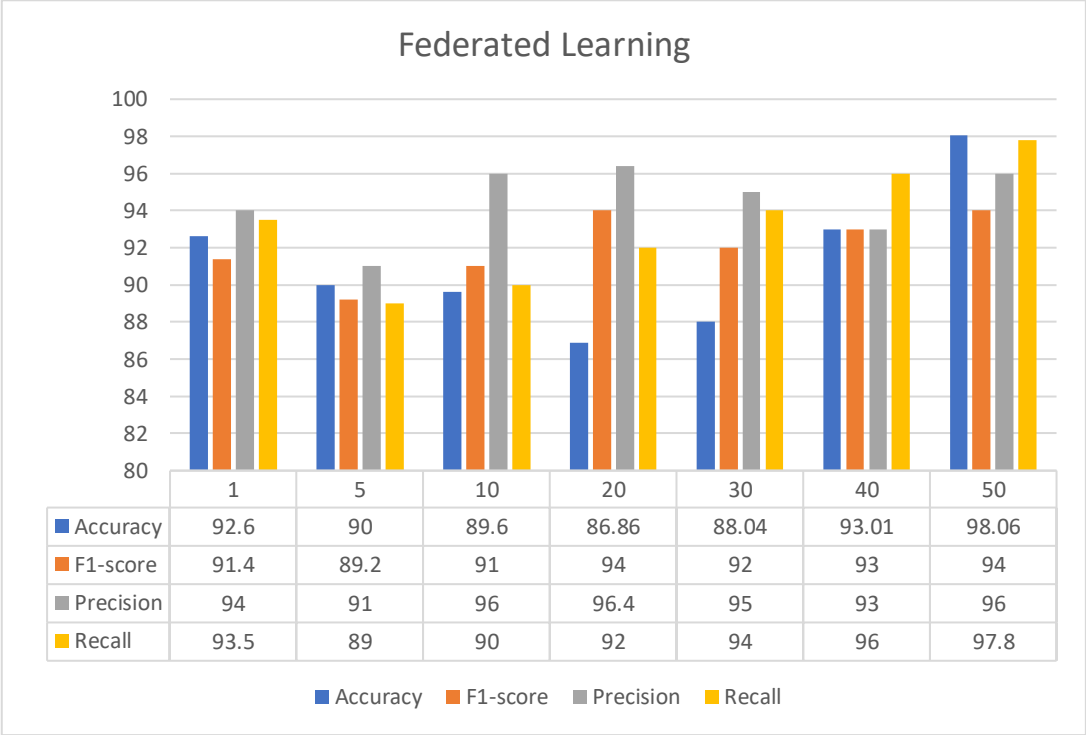


Figure 9 Recall Value comparison between Distributed and Centralized Approach

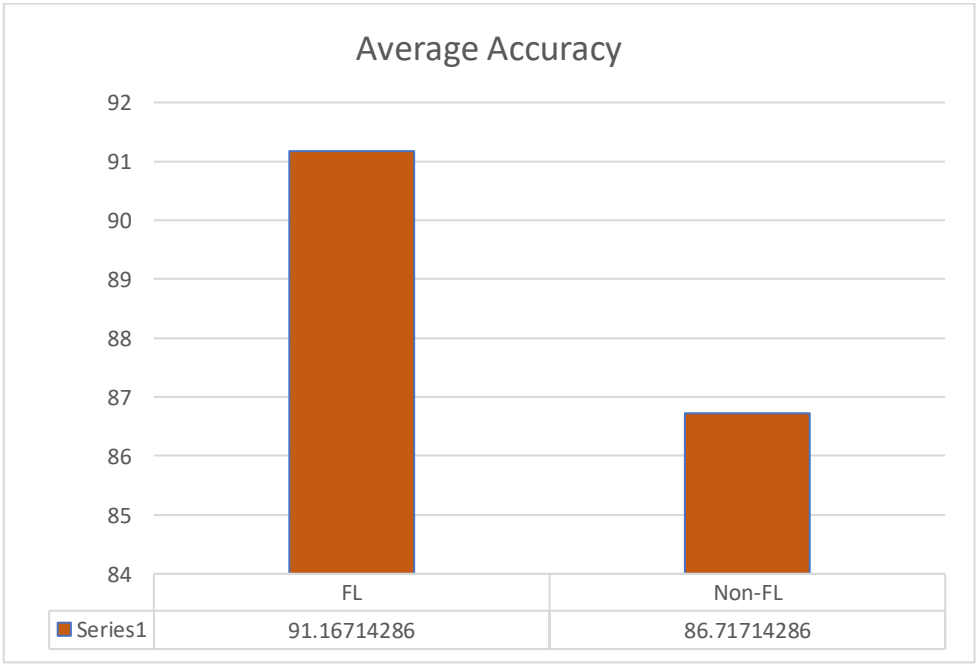
From the figure, it is evident that the recall of both FL and Non-FL increases with the increase in epochs, with the highest recall value of 97.8 and 94.3 for FL and Non-FL respectively at epoch 50.



	1	5	10	20	30	40	50
Accuracy	92.6	90	89.6	86.86	88.04	93.01	98.06
F1-score	91.4	89.2	91	94	92	93	94
Precision	94	91	96	96.4	95	93	96
Recall	93.5	89	90	92	94	96	97.8

Figure 10 Result Comparison between Distributed and Centralized Approach

From the above figure, it is observable that the presentation indicators such as accuracy, recall, F1-score, and accuracy for our suggested method i.e., Federated learning. The accuracy of FL is 98.06, F1-score is 94, and precision value is 96 as well as recall value is 97.8.



	FL	Non-FL
Series1	91.16714286	86.71714286

Figure 10 Average Accuracy Comparison between Distributed and Centralized Approach

The average precision of FL and Non-FL was shown in the above graph. Our proposed FL version, as compared to the non-FL variation, delivers more accuracy with fewer epochs.
This work demonstrates that, in contrast to the conventional DDos detection approach, the combination of deep learning and FL approach can accomplish privacy-protected DDoS categorization in the cloud without losing accuracy.

CONCLUSION

In conclusion, in contrast to the traditional DDoS detection approach, our work shows that using a deep learning and FL approach together may successfully categories DDoS attacks in the cloud while maintaining accuracy. The proposed federated learning-based DDoS detection system has been tested and compared with a non-FL version for the accurate identification and categorization of assaults inside heterogeneous networks. The findings of the comparative testing indicate demonstrates the suggested system performs better than the intrusion detection system without FL techniques in terms of accuracy, F1-score, precision, and recall metrics. Our suggested approach's accuracy, F1-score, precision, and recall were determined to be 98.06, 94, 96, and 97.8 correspondingly. This research exhibits the capability of federated learning in improving the accuracy of anomaly detection systems.
As a Future scope of this study, we intend to construct a new model to prevent or lessen DDoS attacks based on the results of the FL and CNN classification method employed in this research.

REFERENCES

[1] J. Li, L. Lyu, X. Liu, X. Zhang, and X. Lyu, “FLEAM: A Federated Learning Empowered Architecture to Mitigate DDoS in Industrial IoT,” IEEE Trans. Ind. Informatics, vol. 18, no. 6, pp. 4059–4068, 2022, doi: 10.1109/TII.2021.3088938.

[2] G. de Carvalho Bertoli, L. Alves Pereira Junior, O. Saotome, and A. L. dos Santos, “Generalizing intrusion detection for heterogeneous networks: A stacked-unsupervised federated learning approach,” Comput. Secur., vol. 127, no. January, 2023, doi: 10.1016/j.cose.2023.103106.

[3] S. Agrawal et al., “Federated Learning for intrusion detection system: Concepts, challenges and future directions,” Comput. Commun., vol. 195, pp. 346–361, 2022, doi: 10.1016/j.comcom.2022.09.012.

[4] A. K. Shukla and A. Sharma, "Cloud Base Intrusion Detection System using Convolutional and Supervised Machine Learning," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10112007.

[5] A. K. Shukla and A. Sharma, "Classification and Mitigation of DDOS attacks Based on Self-Organizing Map and Support Vector Machine," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10111988.

[6] A. K. Shukla and A. Sharma, "Distributed Attacks Classification Based on Radical Basis Function and Particle Swarm Optimization In Hypervisor Layer," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-4, doi: 10.1109/ISCON57294.2023.10112162.

[7] A. K. Shukla and A. Sharma, "Reduce false intrusion alerts by using PSO feature selection in NSL-KDD dataset," 8th International Conference on Computing in Engineering and Technology (ICCET 2023), Hybrid Conference, Patna, India, 2023, pp. 226-231, doi: 10.1049/icp.2023.1495.

[8] Shukla, Aditya Kumar, and Ashish Sharma. "Cloud Data Security by Hybrid Machine Learning and Cryptosystem Approach." International Journal of Intelligent Systems and Applications in Engineering 12.2s (2024): 01-14.
[9] C. Xu, Y. Qu, Y. Xiang, and L. Gao, "Asynchronous Federated Learning on Heterogeneous Devices: A Survey," pp. 1–49, 2021, [Online]. Available: http://arxiv.org/abs/2109.04269
[10] Petrakopoulos, Vasilis. "DDoS detection using trust-aware federated learning for heterogeneous collaborators." (2022).
[11] A. Cheema, M. Tariq, A. Hafiz, M. M. Khan, F. Ahmad, and M. Anwar, "Prevention Techniques against Distributed Denial of Service Attacks in Heterogeneous Networks: A Systematic Review," Secur. Commun. Networks, vol. 2022, 2022, doi: 10.1155/2022/8379532.
[12] R. Doriguzzi-Corin and D. Siracusa, "FLAD: Adaptive Federated Learning for DDoS Attack Detection," pp. 1–12, 2022, [Online].
[13] Tang, Zhongyun, Haiyang Hu, and Chonghuan Xu. "A federated learning method for network intrusion detection." Concurrency and Computation: Practice and Experience 34.10 (2022): e6812.
[14] J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," Electron., vol. 9, no. 6, pp. 1–21, 2020, doi: 10.3390/electronics9060916.
[15] A. E. Cil, K. Yildiz, and A. Buldu, "Detection of DDoS attacks with feed forward based deep neural network model," Expert Syst. Appl., vol. 169, no. December 2020, p. 114520, 2021, doi: 10.1016/j.eswa.2020.114520.
[16] X. Sáez-de-Cámara, J. L. Flores, C. Arellano, A. Urbieta, and U. Zurutuza, "Clustered Federated Learning Architecture for Network Anomaly Detection in Large Scale Heterogeneous IoT Networks," 2023, [Online].
[17] S. Arisdakessian, O. A. Wahab, A. Mourad, H. Otrok, and M. Guizani, "A Survey on IoT Intrusion Detection: Federated Learning, Game Theory, Social Psychology, and Explainable AI as Future Directions," IEEE Internet Things J., vol. 10, no. 5, pp. 4059–4092, 2023, doi: 10.1109/JIOT.2022.3203249.
[18] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated-Learning-Based Anomaly Detection for IoT Security Attacks," IEEE Internet Things J., vol. 9, no. 4, pp. 2545–2554, 2022, doi: 10.1109/JIOT.2021.3077803.
[19] G. De Carvalho Bertoli, L. A. P. Júnior, and O. Saotome, Improving detection of scanning attacks on heterogeneous networks with Federated Learning, vol. 49, no. 4. Association for Computing Machinery, 2022. doi: 10.1145/3543146.3543172.
[20] N. N. Dao et al., "Securing Heterogeneous IoT With Intelligent DDoS Attack Behavior Learning," IEEE Syst. J., vol. 16, no. 2, pp. 1974–1983, 2022, doi: 10.1109/JSYST.2021.3084199.
[21] Li, Kun, et al. "Distributed network intrusion detection system in satellite-terrestrial integrated networks using federated learning." IEEE Access 8 (2020): 214852-214865.
[22] Dora, V. Raghava Swamy, and V. Naga Lakshmi. "Optimal feature selection with CNN-feature learning for DDoS attack detection using meta-heuristic-based LSTM." International Journal of Intelligent Robotics and Applications 6.2 (2022): 323-349.
[23] Lv, Dingyang, et al. "DDoS Attack Detection Based on CNN and Federated Learning." 2021 Ninth International Conference on Advanced Cloud and Big Data (CBD). IEEE, 2022.