Distributed Denial of Service Attack Prediction over Software Defined Networks Using Federated Learning

Nisarg Mehta (✉ nisarg.mmtcs21@sot.pdpu.ac.in)
Pandit Deendayal Energy University

Madhu Shukla
DEPARTMENT of Computer Engineering Artificial intelligence & Big Data Analytics Marwadi University, India

Pooja Shah
Pandit Deendayal Energy University

Samir Patel
Pandit Deendayal Energy University

Darshit Shah
Pandit Deendayal Energy University

Kishan Makadiya
DEPARTMENT of Computer Engineering Marwadi University

Research Article

Keywords: SDN, DDOS, Federated Learning

Posted Date: April 26th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2832358/v1

Additional Declarations: No competing interests reported.

Distributed Denial of Service Attack Prediction over Software Defined Networks Using Federated Learning

Nisarg Mehta [a], Madhu Shukla [b], Pooja Shah [a], Samir Patel [a], Darshit Shah [a], Kishan Makadiya [c]

a- Computer Science Engineering Department School of Technology Pandit Deendayal Energy University, India

b- DEPARTMENT of Computer Engineering Artificial intelligence & Big Data Analytics Marwadi University, India

c- DEPARTMENT of Computer Engineering Marwadi University, India

Compounding author Nisarg Mehta – nisarg.mmtcs21@sot.pdpu.ac.in, mobile no - +917984611417

Abstract

In today's digital age, Distributed Denial of Service (DDoS) attacks have emerged as a formidable threat to the availability of online services. The ability to predict these attacks in advance is a crucial element in ensuring uninterrupted access to these services. This is where our proposed methodology comes into play. We propose the use of Federated Learning in the prediction of DDoS attacks on Software Defined Networks (SDN). Federated Learning is a cutting-edge approach that allows multiple agents, such as network devices, to collaborate and learn a shared model without the need to share their raw data. Our proposed system leverages the collective intelligence of SDN-enabled network devices to construct a prediction model that can detect DDoS attacks in real time. The experimental results of our study demonstrate the efficacy of our approach in detecting DDoS attacks with a high degree of accuracy and minimal instances of false alarms. We believe that our proposed methodology can prove to be an invaluable tool for service providers in their efforts to prevent DDoS attacks and preserve the availability of online services.

Keywords- SDN, DDOS, Federated Learning

1.Introduction

Software-defined networking (SDN) is a networking method that tries to fix the problems caused by the lack of centralized control on traditional networks. It does this by separating the control plane from the data plane. Information and Communication Technology (ICT) plays a major part in our social and economic life in the digital era and may have a substantial influence on a country's GDP.As technology advances, cyber dangers and attacks increase, rendering previously protected information systems vulnerable. New ICT may also increase network security vulnerabilities. Cyberspace is now a theater of operations in warfare, and the United States Cyber Command has been elevated to the status of unified combatant command, highlighting the necessity to ensure the security of online data. In order to maintain the dependability of work and render hostile cyber warfare ineffectual, it is vital to design a comprehensive plan to defend the ever-changing digital environment.[1] To address the multiple cybersecurity vulnerabilities and threats, a flexible, scalable, and cost-effective solution is required. As the pace of digitalization accelerates quickly, cybersecurity experts and academics throughout the globe are working toward the shared objective of building a safe and trustworthy online environment. According to a survey by Cisco, 30% of global businesses were victims of cybercrime in 2019. DoS and DDoS assaults are the most prevalent forms of cyberattacks on the Internet. In a DDoS assault, attackers control a network of hacked computers (a "botnet") and use them to flood their targets with requests, causing the victim's resource-constrained destination to become overloaded and unable to serve genuine customers.[1] In software-defined networking (SDN), the control plane and data plane are kept distinct. This allows for administration responsibilities like traffic monitoring and data routing to be managed by a centralized software controller. Because it is software-based, this virtualized network may operate independently and provide enhanced flexibility, efficiency, and dependability. In addition to facilitating commercial interaction and data exchange between internal and external customers, SDN may be used to manage a company's essential tasks and secure users' personal information.[12] It is also put to use in the administration of industrial IoT devices, where it encrypts the data collected by sensors and performs a speedy analysis to provide useful insights to enterprises. SDN enables low-cost, high-efficiency IoT administration by enhancing application and analytics performance.

Distributed denial-of-service (DDoS) attacks, which involve flooding a network or server with traffic in an effort to interrupt its regular operations, are a possible danger to SDN systems. Due to its decoupled and centralized control plane, an SDN network might be particularly susceptible to distributed denial of service attacks. Firewalls, intrusion prevention systems, rate limiting, and traffic shaping are all common hardware and software-based techniques used by SDN systems to combat this problem. While these techniques may help, they may not be enough to fend against sophisticated DDoS assaults. When it comes to machine learning, federated learning is a method that allows for the training of models using data that is dispersed over several devices or places. Potentially, it might be used to lessen the effect of Distributed Denial of Service (DDoS) assaults by spreading the load over a larger number of computers or geographically dispersed nodes. Using federated learning, one may also train machine learning models to detect and mitigate DDoS assaults in real time by automatically blocking or rerouting traffic from suspected sources based on a set of predefined rules. The use of federated learning in combination with other security measures is encouraged

Diagram of SDN DDOS ATTACK

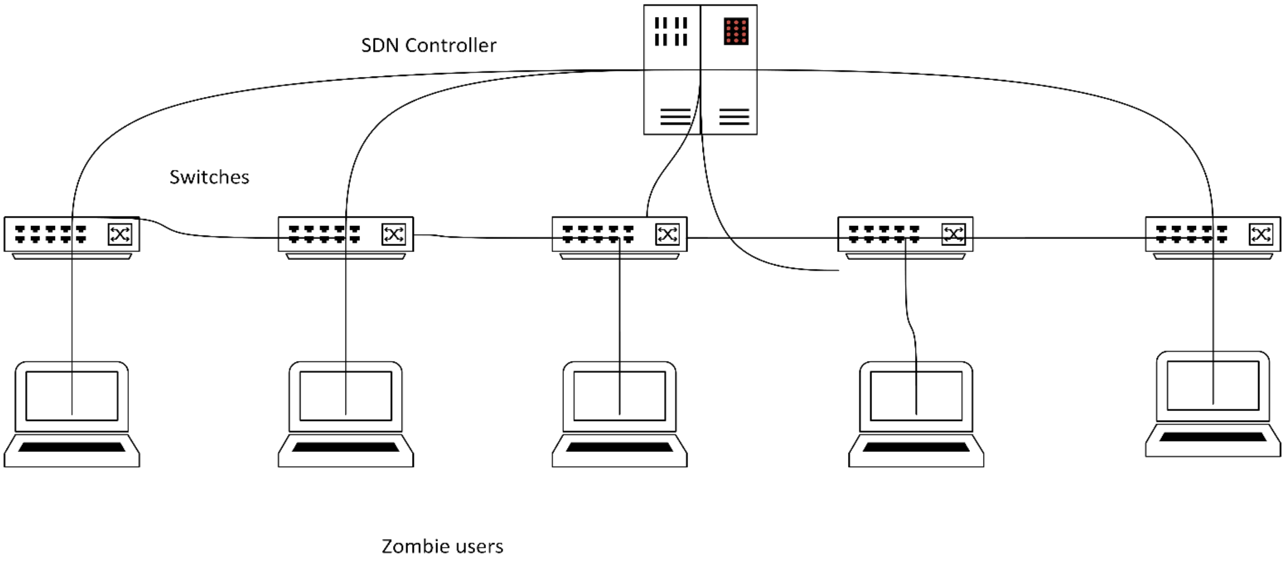


Figure number 1

2 LITERATURE REVIEW

This section elaborates on most relevant state of the art research in the domain of DDoS attack in SDN.

S. Haider *et al 2020*, provided a deep Convolutional Neural Network (CNN) ensemble framework for detecting DDoS assaults in SDN The authors suggest a new method employing deep learning methods to increase the detection rate and decrease the computational cost of DDoS assaults in SDNs. Standard and hybrid versions of the most innovative deep learning techniques were used to evaluate the proposed framework on a flow based SDN dataset. The findings demonstrated that the suggested method successfully reduced the computational complexity of the issue while simultaneously increasing the detection rate. This research also delves into future large-scale distributed networks and the diverse ways for detection and prevention based on deep learning ensembles. In order to ensure the security of the digital era and to address cyber security challenges from a variety of angles, the authors believe that novel and innovative research is required.[2]

In order to reduce the impact of DDoS assaults on Internet Service Provider (ISP) networks, Tuan et al 2020 have developed a strategy that uses machine learning algorithms with SDN. SDN's centralized management and continuous monitoring are put to use in the suggested method to detect and stop DDoS assaults before they create costly interruptions. To prove the efficiency of the suggested strategy in preventing and reducing DDoS assaults in ISP networks, the authors conducted simulated studies.

The possible benefits and drawbacks of utilizing machine learning and SDN to protect ISP networks against DDoS attacks are explored in this article. The suggested technique seeks to provide a more effective and efficient method for detecting and blocking DDoS assaults in ISP networks by capitalizing on the capabilities of these technologies. As a result, the authors believe that the suggested approach has the potential to greatly enhance ISP network security and dependability.[3]

To identify and mitigate DDoS assaults in the industrial IoT (IIoT), Du and Wang 2020 reported a pseudo-honeypot strategy. Pseudo-honeypots, dummy systems that behave like genuine ones to lure inalicious traffic, are deployed with the help of SDN in the suggested approach. In order to identify and mitigate DDoS assaults in IIoT systems, the article explains how SDN may be used to provide centralized management and real-time monitoring of network traffic. To evaluate how well the suggested technique would work in detecting and mitigating DDoS assaults in IIoT systems, the authors ran simulated tests. The findings demonstrated that the suggested technique successfully detected and mitigated DDoS assaults with a low false positive rate and a high detection rate. Both the potential benefits and drawbacks of employing SDN-enabled pseudo-honeypots for DDoS attack detection and prevention in IIoT systems are covered in this study. Authors believe that the suggested approach has the potential to greatly enhance IIoT system security and dependability.[4]

Singh et al 2020. developed a DDoS detection and prevention system that makes use of network function virtualization (NFV) and SDN. In order to identify and prevent DDoS attacks, the suggested system makes use of both NFV and SDN. The former allows the deployment of virtualized network functions (VNFs), while the latter permits centralized management and real-time monitoring of network traffic. This study explores how DDoS assaults may be detected with the use of machine learning techniques and how VNFs can be deployed to counteract these attacks. To gauge how well the suggested system might detect and avert DDoS assaults, the scientists ran simulated trials. With a low false positive rate and a high detection rate, the findings demonstrated that the suggested system successfully identified and prevented DDoS assaults. The study also covers the benefits and drawbacks of using NFV and SDN for DDoS detection and prevention, as well as the difficulties associated with deploying the suggested system in actual networks. The authors sum up by saying that the suggested method has the potential to greatly increase network security and dependability by allowing the quick deployment of VNFs for DDoS detection and prevention.[5]

The application of Artificial Intelligence (AI) to prevent DDoS assaults in SDN is explored by Houda et al 2020 [6]. In order to combat DDoS assaults, the authors of this research suggest using machine learning techniques to detect such attacks and intelligent routing algorithms to reroute traffic elsewhere. The benefits and drawbacks of employing AI for DDoS attack detection and prevention in SDN, as well as the difficulties of applying the suggested technique in real-world networks, are discussed in this article. This research uses simulation results to prove that the suggested method can successfully identify and mitigate DDoS assaults in SDN. With a low false positive rate and a high detection rate, the findings demonstrated that the suggested method successfully identified and prevented DDoS assaults. The advantages of employing AI to add intelligence to SDN-based networks, as well as the scalability and flexibility of the suggested solution, are also discussed by the authors. In conclusion, the authors state that AI has the potential to greatly increase the security and dependability of SDN-based networks by allowing for the early identification and prevention of DDoS assaults.

Nugraha and Murthy 2020 [7] suggested the use of deep learning strategies for the purpose of detecting slow DDoS assaults in SDN. The difficulties of identifying slow DDoS assaults, which are characterized by a low amount of traffic over a lengthy period of time, and the application of deep learning algorithms to detect them are discussed in this work. To detect sluggish DDoS assaults, the authors suggest using CNN to monitor network data. To determine whether or not the suggested method is useful in identifying sluggish DDoS assaults in SDNs, the authors performed simulated tests. The findings revealed that the suggested method has a low false positive rate and a high detection rate when used to identify sluggish DDoS assaults. The authors also address the difficulty of applying the suggested technique in real-world networks, as well as the possible benefits and drawbacks of utilizing deep learning for DDoS attack detection in SDN. The authors believe that deep learning has the potential to greatly increase the security and dependability of SDN-based networks by facilitating the quick identification of sluggish DDoS assaults.

The study conducted by McMahan et al 2017 [8], is groundbreaking because it proposes the idea of federated learning, a method for training machine learning models on distributed data without requiring the data to be centralized. In this research, authors offer a distributed approach to training deep neural networks, in which numerous devices pool their data and computational resources. The difficulty of federated learning arises from the fact that its constituent devices, which may be linked through slow or unreliable connections, must constantly exchange data and model changes. As a solution to this problem, the authors offer a quantization-and-scarification-based approach to compressing model updates, which both decreases the amount of communication needed and preserves the model's correctness. Data augmentation and weight averaging are introduced as methods for coping with the unbalanced nature of decentralized data and increasing the model's generalizability introduced by the authors. This study proposes a novel method for training ML models using distributed data, which has had far-reaching effects in the field and has been the basis for several follow-up studies. Federated learning is a method for training machine learning models on distributed data without the requirement to concentrate the data in a single place, and offers a thorough examination of the area. The many reasons and uses for federated learning, as well as the technological difficulties it presents, are discussed, as are the numerous solutions that have been presented. The balance that must be struck between reduced communication costs and improved model accuracy during federated learning is one of the review's central concerns. Quantization and scarification are only two of the ways the authors explore in order to reduce the amount of conversation needed during training; they also talk about the costs and benefits in terms of model accuracy that come with utilizing these methods. This paper also delves into the particular problems and possibilities presented by the use of federated learning in certain domains, such as natural language processing, computer vision, and healthcare. Overall, gives a thorough and current overview of the topic of federated learning, making it a significant resource for scholars and practitioners in this dynamic domain of machine learning.[9]

Li et al 2022[10] investigates the application of federated learning to the problem of preventing IIoT systems from being compromised by distributed DDoS attacks. The goal of a DDoS attack is to make a network or system inaccessible to its intended audience by flooding it with traffic. In order to lessen the severity of DDoS assaults on IIoT systems, the authors of this study suggest a Federated Learning Empowered Architecture (FLEAM). Federated learning is used in this architecture to train machine learning models using data that has been gathered independently from all of the devices in the IIoT network. By using these models to detect and block DDoS attack traffic, the severity of the attacks may be reduced. As the federated learning models are constantly updated based on fresh data and input from participating devices, the FLEAM architecture is particularly well-suited to the detection and mitigation of DDoS assaults in real time. Additionally, addressing data privacy and security issues in the IIoT environment, federated learning allows the training of models on sensitive or personal data that may not be

References	Conclusion
[2]	The convolutional neural network functions effectively inside the CNN ensemble architecture.
[3]	TCP Flooding attack and Internet Control Message Protocol (ICMP) flood assault in SDN-based ISP networks.
[4]	When using a double honeypot, detection performance is prioritized.
[5]	Different protocols, such as UDP and TCP, may have different thresholds.
[6]	Wisdom SDN has reached the level of complete intelligence for DDoS.
[7]	Through the use of the SDN controller's REST API, the detection module can gather traffic flow statistics from SDN switches and then analyze this data to identify a slow DDoS assault.
[8]	There is a need to increase the performance of solutions, and recent approaches in the literature for identifying DDoS assaults in SDWSN do not necessarily take limited networks into account.
[9]	The traffic resulting from DDoS attacks is analyzed and scrutinized.
[10]	frequency of migration optimized to reduce network resource waste while protecting against attacks
[12]	Superior to state-of-the-art solutions in throughput by twenty-one%
[13]	SDN network and determine the amount of performance loss that may be attributed to a DDoS assault on an SDN network.
[14]	Attacks against software-defined networks that cause a slow denial of service
[15]	There is a need for flexible and dynamic techniques to secure and grow fog-to-things infrastructure, and the possibility for an SDN-based architecture has been suggested.
[16]	By dynamically managing its infrastructure and services, SDN offers a viable solution to networking consumers.
[17]	SDN offers several security-related characteristics.
[18]	SDN control plane decentralization and tuple spaces
[19]	Pushback is launched in the event of a massive volume assault that exceeds the capacity.

Table 2.1 Literature Review

3 Dataset

Through the use of machine learning and deep learning techniques, researchers take advantage of the Mininet emulator's [19] generation of an SDN-specific data set for use in traffic categorization. Initial steps in the project will include configuring a single Ryu controller [20] to oversee a total of ten unique Mininet layouts. A network simulator may mimic both benign and malicious network activity, including TCP, UDP, and ICMP packets (TCP Syn attack, UDP Flood attack, and ICMP assault, respectively). The dataset consists of twenty-three characteristics as mentioned in table 3, some of which were obtained computationally and some directly from the switches. The switch-id, the number of packets, the number of bytes, the time in seconds, the time in nanoseconds, the source IP address, the destination IP address, and the time in seconds and nanoseconds are all characteristics that may be obtained. The quantity of data received by the switch port is shown by its byte count, whereas the number of bytes sent is represented by the port number. The current time and date are shown numerically in the dt field, which is updated every 30 seconds. The Port Bandwidth is equal to the product of the transmit (TX kbps) and receive (RX kbps) data rates. The data specifications are mentioned in table 2. The simulation is performed for another set amount of time, and maybe more data is gleaned.[11]

Table 2 Dataset specification table

Flow monitoring interval	30 sec
Number of classes	2
Class label 0	Benign traffic
Class label 1	malicious traffic
Network simulation is run	250 minutes
Total Data collected	1,04,345

Table 3 SDN DDOS Dataset table

Extracted Features	**Calculated Features**
Switch-id	Packet per flow which is the packet count during a single flow
Packet_count	Byte per flow is the byte count during a single flow
byte_count	The packet Rate is number of packets send per second and calculated by dividing the packet per flow by monitoring interval
duration_sec	number of Packet_ins messages
duration_nsec which is duration in nano-seconds	total flow entries in the switch
total duration is sum of duration_sec and durstaion_nsec	tx_kbps
Source IP	rx_kbps
Destination IP	data transfer

Port number	receiving rate
tx_bytes is the number of bytes transferred from the switch port	Port Bandwidth is the sum of tx_kbps and rx_kbps
rx_bytes is the number of bytes received on the switch port	
dt field show the date and time which has been converted into number	

4 Methodology

Federated Learning(FL)

Federated Learning(FL)is a machine learning methodology that can be used to train a single model on several nodes or devices without exchanging raw data. Each node in a federated learning network autonomously trains its own model with its own data and then shares the improved model with a master node. The server compiles and merges all new model versions into the master, global model. Once the global model has been modified, it is sent back to the devices so they may use it to enhance their own internal representations. This procedure is repeated until the global model achieves an acceptable level of accuracy. The most significant benefit of federated learning is that it removes the need to centralize data, enabling machine learning models to be trained utilizing any accessible data. This allows models to be trained without compromising the confidentiality of sensitive or secret data, which is particularly useful in circumstances when the data is sensitive or confidential . The concept of FL can be visualized as in the algorithm and the description of terminology regarding the same followed by Figure number 2.

```
    Algorithm 1 – Federated average. The j client are index by p; G is local minimum batch size E is local
number of epoch and R is learning rate.
    Server executes.
    Initialize w0 For each round t= 1,2,3,4,5,6,7,8,9
    "j←max (F ·j,1)"
    Ht←(random set of j clients)
    For each client j ∈Ht in paralle do
    Wij+1 ←Client update( j ,Wi)
    $W_{i+1}$  $W_i^j$+1
    Client update (j,W): ‖ run on client j
    G←1(split Rj into Batch size of G)
    For each local epoch E from 1 to 46 do
    For batch g ∈G do W←w-m∇U(W;G)  Return W to the sever
```

4.1 The FL Algorithm terminology

W_i-model weights on communication round #i

W_i^j - model weights on communication round #i on client #j

A- Fraction of client performing computation on each round

D- Number of training pass each client makes over dataset on each round.

G- The local mini batch size used for client updates

m – learning rate

Q_j- set of data point on client j

R_j – Number of data point on client j

E_i(W)- Loss U (A_x,$B_{x,}$C) with parameters W

p – indexing

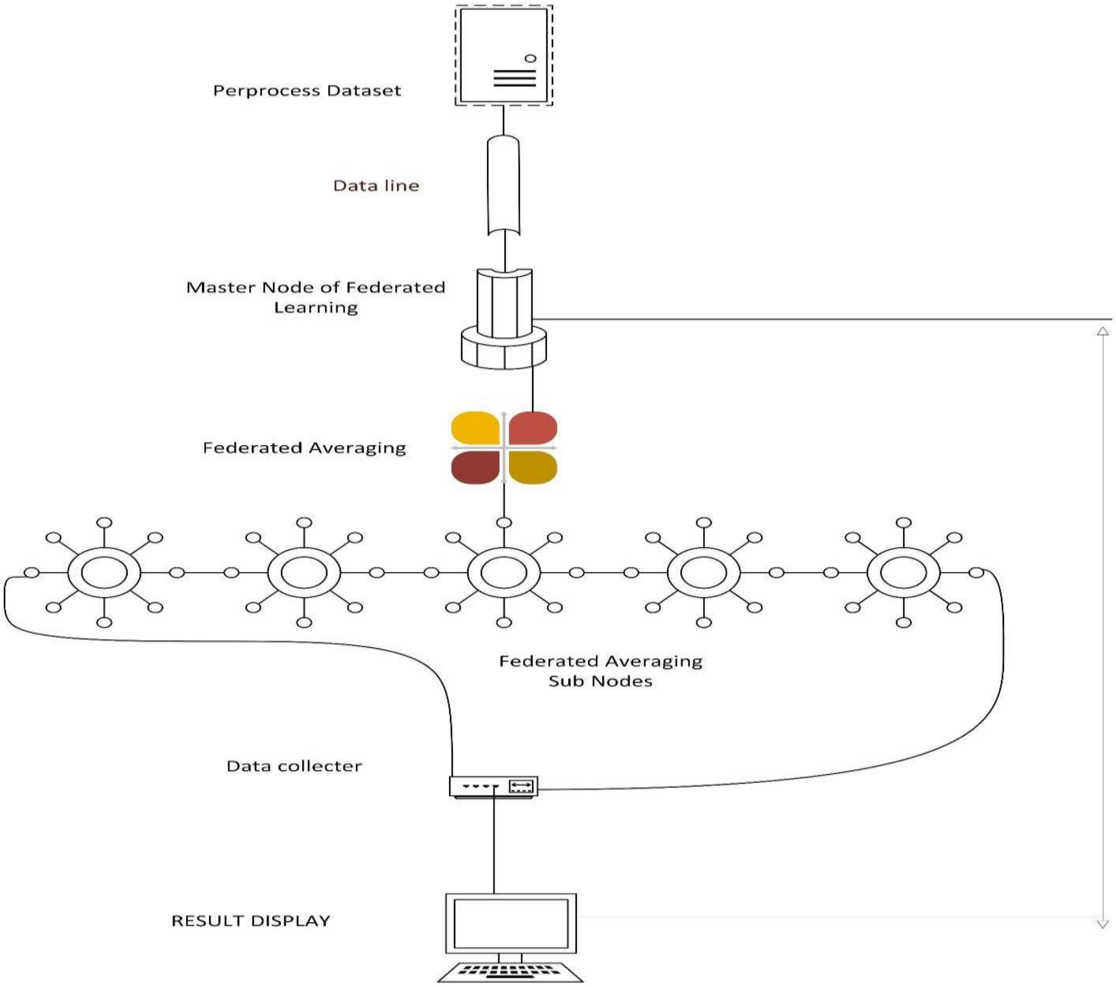
Equation

$Fj\ (W)\ = \frac{1}{nj} \sum_{j \in Q_j} \quad E_j(W)$ (1)

$gj\ =\ \nabla Fj\ (Wj)$

This corresponds to a full-batch (non-stochastic) gradient descent. For the current global model W^i, the average gradient on its global model is calculated for each client *J*.

F- Fraction of clients participating in that round **T** - No. of training passes each client makes over its local dataset each round **G** - Local minibatch size used for client updates



SDN DDOS Federated Learning algorithm working Figure number2

5 Experimental Setup

Table 4 The experimental setup is described in the table below:

Platform	● Google Collab
Capabilities	● Tesla T-4 GPU ● 16 GB GGDR6 ● DISK 78.19 GB
Federated Learning	Master Slave Model
Data Set	described in section
hyperparameters	● NO of rounds -9 ● Epoch -45

	- Bach size -10 - Weight scaling factor - Weigh federated average. - Q Federated average - Learning rate -10 - Comms round -10 - Optimizer SGD
Evaluation Matrix	- Accuracy = $\frac{Number\ of\ \ corract\ perdication}{Toltal\ Perdication}$ - Precision = True Positives / (True Positives + FalsePositives) - Recall = TruePositives / (TruePositives + FalseNegatives) - F1 Score= (2 * Precision * Recall) / (Precision + Recall)

Table: Experimental Setup

6.Results and Discussion

In this study, we aimed to address the issue of DDoS attacks in SDN utilizing the capabilities of federated learning. DDoS attacks have become a major concern for SDN as they can disrupt network functionality and cause significant damage. In order to mitigate these attacks, accurate and timely prediction of them is crucial.

To achieve this goal, we proposed a federated learning approach that utilizes multiple SDN controllers as clients to train a global model while maintaining the privacy of the data. Federated learning is a machine learning approach that allows multiple devices or clients to train a model without sharing their data with a central server. This approach is particularly useful in the case of SDN, as it allows the controllers to share information and learn from one another while preserving the privacy of their data.

The experimental results mentioned in table 5 of the proposed approach demonstrated its effectiveness in predicting DDoS attacks with increasing epoch. The results showed that the proposed approach achieved a high accuracy of 99.39% in identifying DDoS attacks. This high level of accuracy is crucial for the timely detection of DDoS attacks and the implementation of countermeasures.

When it comes to evaluating the performance of a federated learning model, there are a number of metrics that can be used, including accuracy, precision, recall, and F1 score. Each of these metrics provides a different perspective on the model's performance and can be used in combination to gain a more comprehensive understanding of the model's capabilities.
Accuracy is a simple metric that measures the proportion of correct predictions made by the model. It is calculated by dividing the number of correct predictions by the total number of predictions. While accuracy can be a useful metric, it can be misleading if the dataset is imbalanced, meaning that there is a disproportionate number of samples in one class compared to the other.

Precision, on the other hand, is a measure of the model's ability to make correct positive predictions. It is calculated by dividing the number of true positive predictions by the total number of positive predictions made by the model. A high precision indicates that the model is confident in its positive predictions, but it may miss some actual positive examples.

Recall is a measure of the model's ability to identify all of the positive examples in the dataset. It is calculated by dividing the number of true positive predictions by the total number of actual positive examples. A high recall indicates that the model is able to find most of the positive examples, but it may also include some false positive predictions.

F1 Score is the harmonic means of precision and recall. It tries to balance the trade-off between precision and recall and gives more weight to the lower value. It is calculated by 2*(Precision*Recall)/(Precision+Recall). Each of these metrics provides a different perspective on the model's performance and can be used in combination to gain a more comprehensive understanding of the model's capabilities. It's important to choose the right

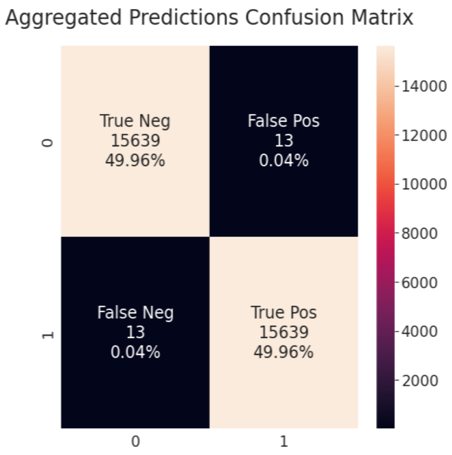
metric depending on the specific use case of the model, as different scenarios have different prioritie
.

The results indicated that the proposed approach is robust and adaptable, as it was able to maintain its high level of accuracy even under varying conditions.

As an extension to this work, we also evaluated the proposed approach under different scenarios, such as varying numbers of clients and different attack types.

Table 5 : Experimental results of federated learning in the prediction of DDOS attack in SDN

Round	accuracy	recall	precision	F1 score
1	98.90%	98.90%	98.90%	98.90%
2	98.83%	98.83%	98.83%	98.83%
3	99.32%	99.32%	99.32%	99.32%
4	99.08%	99.07%	99.08%	99.07%
5	99.39%	99.39%	99.39%	99.39%
6	99.33%	99.33%	99.33%	99.33%
7	99.17%	99.17%	99.17%	99.17%
8	99.54%	99.54%	99.54%	99.54%
9	99.39%	99.39%	99.39%	99.39%



6.Conclusion

In conclusion, the proposed federated learning approach for DDoS attack prediction in SDN showed promising results in terms of accuracy and efficiency. The approach was able to effectively predict DDoS attacks while maintaining the privacy of the data and could serve as a valuable tool for securing SDN networks against such threats. Future work could include extending the proposed approach to other types of network attacks and incorporating other techniques to enhance performance**.**

Declarations

Conflict of Interest: The authors declare that they have no conflict of interest.

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript

.

Ethical Approval

not applicable

Authors' contributions

The author contributions for the manuscript titled "Distributed Denial of Service Attack Prediction over Software Defined Networks Using Federated Learning" are as follows: Nisarg Mehta was the main researcher for this project and played a significant role in designing the experiments, collecting the data, analyzing the results, and writing the manuscript. Dr. Pooja Shah provided her expertise in the field of network security and SDNs, reviewed the manuscript, and contributed significantly to improving the quality of the paper. Dr. Madhu Shukla reviewed the manuscript and provided her valuable feedback, especially in the area of machine learning and federated learning. Dr. Samir Patel provided his expertise in the field of software engineering and SDN architecture and helped in refining the research methodology and the technical aspects of the manuscript. Mr. Kishan Makdiya played an essential role in data collection and preprocessing, and he was also involved in the implementation and testing of the proposed methodology. Darshit Shah contributed to the literature review, helped in data collection, and assisted in the analysis of the results. All authors have reviewed and approved the final manuscript.

Funding

This research was conducted with no external funding. The authors contributed their time, expertise, and resources to conduct the study, including the acquisition and analysis of data, development of the distributed denial of service attack prediction model, and preparation of the manuscript. The authors have no financial or other conflicts of interest related to this research.

Availability of data and materials

The dataset and materials used in this study are publicly available at https://data.mendeley.com/datasets/jxpfjc64kr/1. The dataset contains network traffic data used in the development and evaluation of the distributed denial of service attack prediction model described in this manuscript. The dataset is accompanied by a data dictionary that provides information on the variables and their definitions. The dataset is available under a Creative Commons Attribution 4.0 International License, which allows for sharing and adaptation of the dataset, provided appropriate credit is given to the authors. The DOI for the dataset is 10.17632/jxpfjc64kr.1.

References

[1] Mehta, N., Sanghavi, P., Paliwal, M., Shukla, M. (2022). A Comprehensive Study on Cyber Legislation in G20 Countries. In: Rajagopal, S., Faruki, P., Popat, K. (eds) Advancements in Smart Computing and Information Security. ASCIS 2022. Communications in Computer and Information Science, vol 1760. Springer, Cham. https://doi.org/10.1007/978-3-031-23095-0_1

[2] S. Haider *et al.*, "A Deep CNN Ensemble Framework for Efficient DDoS Attack Detection in Software Defined Networks," *IEEE Access*, vol. 8, pp. 53972–53983, 2020, doi: 10.1109/ACCESS.2020.2976908.

[3] N. N. Tuan, P. H. Hung, N. D. Nghia, N. van Tho, T. van Phan, and N. H. Thanh, "A DDoS attack mitigation scheme in ISP networks using machine learning based on SDN," *Electronics (Switzerland)*, vol. 9, no. 3, pp. 1–19, 2020, doi: 10.3390/electronics9030413.

[4] M. Du and K. Wang, “An SDN-Enabled pseudo-honeypot strategy for distributed denial of service attacks in industrial internet of things,” *IEEE Trans Industr Inform*, vol. 16, no. 1, pp. 648–657, 2020, doi: 10.1109/TII.2019.2917912.

[5] A. K. Singh, R. K. Jaiswal, K. Abdukodir, and A. Muthanna, “ARDefense: DDoS detection and prevention using NFV and SDN,” *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, vol. 2020-Octob, pp. 236–241, 2020, doi: 10.1109/ICUMT51630.2020.9222443.

[6] Z. Abou El Houda, L. Khoukhi, and A. Senhaji Hafid, “Bringing Intelligence to Software Defined Networks: Mitigating DDoS Attacks,” *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2523–2535, 2020, doi: 10.1109/TNSM.2020.3014870.

[7] B. Nugraha and R. N. Murthy, “Deep Learning-based Slow DDoS Attack Detection in SDN-based Networks,” *2020 IEEE Conference on Network Function Virtualization and Software Defined Networks, NFV-SDN 2020 - Proceedings*, pp. 51–56, 2020, doi: 10.1109/NFV-SDN50289.2020.9289894.

[8] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-Efficient Learning of Deep Networks from Decentralized Data.” PMLR, pp. 1273–1282, Apr. 10, 2017. Accessed: Dec. 27, 2022. [Online]. Available: https://proceedings.mlr.press/v54/mcmahan17a.html

[9] Q. Yang, Y. Liu, Y. Cheng, Y. Kang, T. Chen, and H. Yu, “Federated Learning,” *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 13, no. 3, pp. 1–207, Dec. 2019, doi: 10.2200/S00960ED2V01Y201910AIM043.

[10] 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 Industr Inform*, vol. 18, no. 6, pp. 4059–4068, Jun. 2022, doi: 10.1109/TII.2021.3088938.

[11] D. M. G. S. N. Ahuja, “dataset_sdn.” doi: 10.17632/jxpfjc64kr.1.

[12] Chen, J.I.Z. and Smys, S., 2020. Social multimedia security and suspicious activity detection in SDN using hybrid deep learning technique. *Journal of Information Technology*, *2*(02), pp.108-115.

[13] T. A. Pascoal, I. E. Fonseca, and V. Nigam, “Slow denial-of-service attacks on software defined networks,” *Computer Networks*, vol. 173, no. October 2019, 2020, doi: 10.1016/j.comnet.2020.107223.

[14] P. Krishnan, S. Duttagupta, and K. Achuthan, “SDN/NFV security framework for fog-to-things computing infrastructure,” *Softw Pract Exp*, vol. 50, no. 5, pp. 757–800, 2020, doi: 10.1002/spe.2761.

[15] R. Swami, M. Dave, and V. Ranga, “Voting-based intrusion detection framework for securing software-defined networks,” *Concurr Comput*, vol. 32, no. 24, pp. 1–16, 2020, doi: 10.1002/cpe.5927.

[16] A. M. Abdelrahman *et al.*, “Software-defined networking security for private data center networks and clouds: Vulnerabilities, attacks, countermeasures, and solutions,” *International Journal of Communication Systems*, vol. 34, no. 4, pp. 1–20, 2021, doi: 10.1002/dac.4706.

[17] S. Belguith, M. R. Asghar, S. Wang, K. Gomez, and G. Russello, “SMART: Shared memory based SDN architecture to resist DDoS ATtacks,” *ICETE 2020 - Proceedings of the 17th International Joint Conference on e-Business and Telecommunications*, vol. 3, no. Icete, pp. 608–617, 2020, doi: 10.5220/0009864906080617.

[18] M. Fischer, “SDN / NFV-based DDoS Mitigation via Pushback,” 2020.

[19] M. Paliwal and K. K. Nagwanshi, “Effective Flow Table Space Management Using Policy-Based Routing Approach in Hybrid SDN Network,” *IEEE Access*, vol. 10, pp. 59806–59820, 2022, doi: 10.1109/ACCESS.2022.3180333.

[20] Contributors, M.P. *Mininet*. Available at: http://mininet.org/ (Accessed: January 25, 2023).

[21] *Ryu SDN Framework*. Available at: https://ryu-sdn.org/ (Accessed: January 25, 2023).