An analysis on Federated Learning Approaches for Scalable DDoS Protection in Multi-Controller SDNs

Gunjani J. Vaghela[1], Simran A. Sodha[2], and Punit C. Trivedi[3]
[1]*Assistant Professor, Atmiya University, Rajkot, Gujrat, India*
[2,3] *Lecturer, Atmiya University, Rajkot, Gujrat, India*

Abstract*— Software-Defined Networking (SDN) with multi-controller architectures faces significant challenges in detecting Distributed Denial of Service (DDoS) attacks due to the decentralized nature of network control and the volume of data traffic. Traditional approaches struggle to scale efficiently while maintaining high security and privacy standards. This survey paper explores the use of Federated Learning (FL) as an innovative solution for adaptive DDoS attack detection in SDN environments. We review the integration of FL with machine learning (ML) and deep learning (DL) models, highlighting their potential to provide decentralized, privacy-preserving detection mechanisms. Additionally, we examine existing methodologies, evaluate their strengths and weaknesses, and discuss the future research directions, such as improving model accuracy, overcoming communication overhead, and addressing security challenges in large-scale SDN deployments. By combining FL with SDN, this approach promises to enhance DDoS detection systems' scalability and efficiency, offering a robust solution for modern network infrastructures.

Index Terms*— Adaptive DDoS Detection, Federated Learning (FL), Multi-Controller Architectures, Software-Defined Networking (SDN)

I. INTRODUCTION

Networking has become an integral part of our daily lives, transforming how we communicate, collaborate, and access resources. It is essential not only for personal communication but also for business operations and global connectivity. Traditional network architectures, which act as global communication pathways, are now experiencing increased traffic due to the growing demands of emerging technologies [1] With the rise of mobile, cloud, and edge computing, network traffic patterns have significantly changed. Mobile devices have made it possible for users to access the internet from anywhere at any time, introducing diverse sources of traffic. Meanwhile, the expansion of cloud computing and edge technologies has led to the emergence of "east-west" traffic flows, contrasting with traditional "north-south" traffic, creating more complex and passive traffic across wide-area networks [1].Traditional network architectures, which rely on conventional routers and switches, face significant challenges in keeping up with evolving traffic demands. One major issue is the manual configuration of individual devices through vendor-specific interfaces, a time-consuming process that hinders quick responses to dynamic traffic changes. Additionally, traditional networks often struggle with scalability, making it difficult to meet the growing needs of various service providers.To address these limitations, Software-Defined Networking (SDN) has emerged as a flexible and scalable solution. By separating the control plane from the data plane and centralizing network management, SDN enables rapid policy changes and efficient reconfiguration to handle fluctuating traffic patterns [2].

A Distributed Denial of Service (DDoS) attack is a malicious attempt to disrupt the normal functioning of a server, service, or network by overwhelming it with a flood of internet traffic. This is achieved using multiple compromised systems, such as computers or IoT devices that act as a botnet to send a massive volume of requests to the target, making it inaccessible to legitimate users[4].DDoS attacks significantly impact Software-Defined Networking (SDN) due to its centralized control architecture. The SDN controller, as the brain of the network, is a critical target. Overwhelming it with attack traffic can paralyze its ability to manage the data plane, leading to severe service disruptions such as high latency, packet loss, or total network outages. Additionally, attack traffic depletes resources like CPU and memory in both controllers and switches, weakening their ability to handle legitimate traffic. These vulnerabilities amplify scalability issues, making robust detection and mitigation strategies essential [3].The increasing sophistication of DDoS attacks in SDN highlights the need for advanced detection and mitigation strategies that can operate in real-time and adapt to evolving threats. Machine

learning (ML) has emerged as a powerful tool for identifying and mitigating such attacks, leveraging data-driven techniques to detect anomalous traffic patterns with high accuracy.Machine learning (ML) approaches are traditionally centralized, relying on a central server to aggregate data from multiple clients and train a model. The effectiveness of these centralized systems heavily depends on the quality and quantity of the collected training data. This centralized nature can lead to bottlenecks, increased risks to data privacy, and challenges in managing diverse datasets across distributed environments, which are crucial for developing robust and accurate models [4].However, centralized training data is not always accessible, as clients may be unwilling to share their data due to privacy concerns. For example, healthcare providers, patients, insurance companies, and pharmacies often refrain from contributing sensitive data for model training. Similarly, banks, law enforcement agencies, and retail stores may resist centralizing their data for tasks like behavior prediction or recommendation systems. This reluctance creates challenges in gathering sufficient, high-quality training data, which limits the advancement of centralized machine learning approaches. Addressing these concerns is critical for future developments.Federated Learning (FL) is a distributed machine learning method that eliminates the need for centralized data and model training. In this approach, a central server coordinates with multiple clients. The process begins with the central server sending a training task and the current global model to the clients. Each client updates the global model using their local data and returns the updated model to the server. The server aggregates these updates to improve the global model. This process repeats iteratively until a termination condition is met [5].Federated Learning (FL) offers several advantages: (1) it significantly enhances data security and privacy by eliminating the need for data sharing and centralized storage; (2) central servers conserve memory since they don’t store large datasets; and (3) the computational load on central servers is reduced as clients handle parts of the training process. Additionally, FL can deliver high-quality models even when individual client datasets are limited in size, enabling effective collaboration across distributed environments [6].

II. BACKGROUND

Traditional networking architectures face challenges in scalability, flexibility, and centralized management, limiting their ability to meet the demands of modern applications and dynamic traffic patterns. To address these issues, Software-Defined Networking (SDN) emerged as an innovative framework that decouples the control plane from the data plane, enabling centralized control of network behavior. Application Plane**:** This layer consists of network applications and services, such as traffic engineering, security, and monitoring. It defines high-level policies and interacts with the control plane through APIs to dictate how the network should behave.Control Plane: The control plane serves as the brain of the SDN architecture. It manages the decision-making process for routing, traffic handling, and policies. The SDN controller, which resides here, communicates with both the application plane (northbound APIs) and the data plane (southbound APIs).Data Plane: The data plane, or infrastructure layer, is responsible for forwarding packets based on instructions from the control plane. This layer consists of network devices such as switches and routers that follow the control plane's directives without making independent decisions [7]. The separation of the control and data planes, combined with the centralized nature of the control plane, makes SDN particularly susceptible to security threats like Distributed Denial of Service (DDoS) attacks. DDoS attacks are maliciously designed to disrupt normal network operations by overwhelming systems with an excessive volume of connection requests within a short period. This flood of traffic strains the target's resources, leading to slow performance, system crashes, or complete unavailability. The primary goal of such attacks is to deny legitimate users access to critical services by saturating network devices and exhausting their capabilities. By disrupting normal operations, DDoS attacks cause significant service interruptions, undermining the reliability and availability of the targeted systems.Volumetric Attacks: These attacks flood the network with massive amounts of traffic to consume the target's bandwidth. Examples include UDP floods and DNS amplification attacks, which exploit the sheer volume to overwhelm the system.Protocol Attacks: Targeting specific weaknesses in network protocols, these attacks consume server resources or intermediary devices. Examples include SYN floods, Ping of Death, and Smurf attacks, which exploit flaws in the TCP/IP stack. Application Layer Attacks: These attacks focus on the application layer, targeting specific services like HTTP, DNS, or SMTP. Examples

include HTTP GET/POST floods and Slowloris, designed to exhaust resources by mimicking legitimate user behavior [8]. DDoS attacks target the availability of systems by overwhelming them with excessive traffic. Detecting these attacks in real time is challenging due to the evolving nature of attack patterns. Federated Learning offers a solution by enabling multiple SDN controllers to collaboratively train a machine learning model for DDoS detection without sharing sensitive traffic data. FL allows the system to continuously adapt to new attack strategies while preserving privacy, making it an ideal approach for improving the security and resilience of SDN networks against DDoS attacks. Federated Learning (FL) is a distributed machine learning approach that enables the training of a shared model while keeping data decentralized. Clients (devices, nodes, or SDN controllers) perform local training on their data, ensuring that sensitive data remains private. After local training, they send model updates (weights and gradients) rather than the raw data to a central server, which aggregates the updates to improve the global model [9].

A. Federated Learning Architecture:

Global Model: This is the central model maintained by the server that is updated based on aggregated local updates. It is the model that aims to improve over multiple training rounds using contributions from all clients.

Local Updates: Each client trains its own model using its local dataset and sends the model updates (e.g., model weights or gradients) to the server. This allows the central model to evolve without exposing private data.

Server: The central server aggregates the updates from each client using an aggregation function (e.g., Federated Averaging). The global model is then updated with the aggregated results [10].

III. LITERATURE REVIEW

Detecting DDoS attacks in Software-Defined Networks (SDNs) has been an area of active research, especially in the context of multi-controller architectures. Various detection methods have been proposed, ranging from traditional machine learning algorithms to more advanced techniques like federated learning. These approaches aim to address the challenges posed by the dynamic nature of DDoS attacks, such as low-rate or evolving attack patterns. This section presents a comparison of several research efforts that focus on DDoS detection in SDN environments, highlighting the detection methods, types of DDoS attacks targeted, datasets used, and key limitations. The survey by Tao Hu et al. [11] focuses on multi-controller based Software-Defined Networking (SDN), which separates control and data planes to enhance network flexibility and programmability. The paper examines the architecture, benefits, and challenges of multi-controller SDN, including scalability, fault tolerance, and ensuring consistency between controllers. It also reviews various research techniques and approaches used in this area. The paper offers a comprehensive overview of the advantages and challenges of multi-controller SDN, providing valuable insights into its potential applications and limitations. In their paper, Aslan et al. [12] employed clustering techniques, including sequential and incremental K-means, to develop a tunable consistency model for managing network dynamics in distributed SDN controllers. This model adapts to changes in the network and was evaluated through simulations. The results demonstrated its effectiveness in maintaining consistency while addressing the challenges posed by dynamic network conditions, offering a practical solution for consistency management in SDN environments. In their paper, Yuan Zhang et al. [13]. discuss the challenges of reactive fault management in SDN, which can lead to high computational burdens and prolonged recovery times for controllers. To address these issues, they propose a combination of preventive and recovery mechanisms Preventive methods, such as route backups and common fault message storage, can help achieve quicker recovery times, while recovery mechanisms, though effective, may introduce additional computational overhead and longer recovery durations. In their study, S. Mukherjee et al. [14] propose the Load-Constrained Control (LCC) solution, which dynamically adjusts traffic distribution among controllers by periodically monitoring load windows. When the load changes, LCC adjusts the controller pool size to meet current demand. If the load exceeds the maximum capacity of the existing controllers, the system adds new controllers to maintain availability and ensure efficient network operation.

IV. COMPARISON TABLE

Paper Name	Detection Method	DDoS Type	Method	Dataset	Limitations
"DDoS Detection in SDN Using Machine Learning" (2022)	Supervised Learning	Volumetric, Application Layer	Decision Trees, SVM	CICIDS 2017	High computational cost, low scalability in large networks
"Federated Learning for DDoS Detection in SDN" (2023)	Federated Learning	Multi-vector, Protocol	Federated Averaging	Custom SDN Traffic Data	High communication cost, limited model updates
"Real-time DDoS Attack Detection in SDN" (2021)	Deep Learning	Protocol-based	CNN, RNN	ISCX 2016	High false positives, long training time
"A Hybrid Machine Learning Approach for DDoS Detection" (2021)	Ensemble Learning	Volumetric	Random Forest, SVM	KDD99, UNSW-NB15	Requires large labeled data, complex tuning
"DDoS Detection Using Federated Learning" (2024)	Federated Learning	Application Layer	Federated Averaging	Custom Dataset	High client computation costs, data imbalance

IV. INTEGRATING FEDERATED LEARNING INTO SDN CONTROLLERS FOR DISTRIBUTED DDOS ATTACK DETECTION

Federated Learning (FL) can be integrated into SDN controllers for collaborative DDoS detection by enabling multiple controllers to train models without sharing sensitive data. Local Training: Each SDN controller collects local network traffic data to train a model. This data can include packet flow information, traffic patterns, and other relevant metrics. The controller trains a local model on this data to detect potential DDoS anomalies. Model Sharing: Instead of sharing raw data, the controllers share updates to the locally trained models with a central aggregator. These updates contain the learned parameters, such as weights and biases, that capture attack patterns specific to each controller's environment. Model Aggregation: The central server aggregates the local models, combining the updates into a global model. The aggregation process typically uses algorithms like Federated Averaging to combine the learned parameters and create a more generalized model that incorporates diverse data from different controllers. Global Model Deployment: Once the global model is updated, it is sent back to the SDN controllers. These controllers then apply the global model to detect and mitigate DDoS attacks based on the aggregated knowledge from all participating controllers. Iterative Learning: The process is iterative, with controllers continuously training on their local data and exchanging updates with the central server. Over time, the global model improves, becoming more effective at identifying and responding to DDoS attacks in a collaborative and distributed manner. Security and Privacy Challenges Security Challenges in Federated Learning: Model Poisoning Attacks: One of the main security risks in FL is model poisoning, where malicious participants (clients) send manipulated model updates to the central server. These corrupted updates can significantly degrade the performance of the global model, leading to poor DDoS detection accuracy. Inference Attacks: In addition to eavesdropping, adversaries can perform inference attacks where they try to infer sensitive information about the local data used by the clients based on the shared model updates. Future Directions: Edge Computing and FL Integration: With the rise of edge Computing and FL Integration: With the rise of edge computing, integrating FL with edge devices could reduce the computational load on SDN controllers, further improving the scalability and efficiency of DDoS detection systems. This could be particularly beneficial in highly distributed networks where edge devices are closer to the traffic source.

V. CONCLUSION

This survey highlights the significant potential of federated learning (FL)-based methods for DDoS attack detection in multi-controller SDNs, showcasing their ability to enhance data privacy, reduce communication overhead, and provide scalable solutions. The adaptive nature of FL in SDNs allows for real-time, decentralized attack detection, making it an ideal fit for dynamic and distributed network environments. However, challenges remain in areas such as adversarial

resilience and handling non-IID data, calling for further research into hybrid FL-SDN architectures and more robust models to address these issues and improve overall system security.

REFERENCES

[1] Cui, Y., Qian, Q., Guo, C., Shen, G., Tian, Y., Xing, H., & Yan, L. (2021). Towards DDoS detection mechanisms in software-defined networking. Journal of Network and Computer Applications(2021), vol-190, PP-103156, jul 2021.

[2] Jain, S., Kumar, A., Mandal, S., Ong, J., Poutievski, L., Singh, A.,& Vahdat, A. (2013). B4: Experience with a globally-deployed software defined WAN. ACM SIGCOMM Computer Communication Review, (2013) , vol 43 PP- 14, Aug 2013

[3] Vaghela, G., Sanghani, N., & Borisaniya, B. (2024). A Review on DDoS Attack in Controller Environment of Software Defined Network. EAI Endorsed Transactions on Scalable Information Systems (2024), vol-11 PP- 17 jul-2024

[4] Boutaba, R., Salahuddin, M. A., Limam, N., Ayoubi, S., Shahriar, N., Estrada-Solano, F., & Caicedo, O. M. (2018). A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. Journal of Internet Services and Applications (2018), vol- 9 PP- 99 jun 2018.

[5] Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated learning for healthcare informatics. Journal of healthcare informatics research (2021), vol 5 PP-19 Nov 2021.

[6] Ma, X., Liao, L., Li, Z., Lai, R. X., & Zhang, M. Applying federated learning in software-defined networks: A survey. Symmetry (2022),vol-14 PP 195 jan 2022

[7] Mahar, I. A., Libing, W., Maher, Z. A., & Rahu, G. A. (2024, January). A Comprehensive Survey of Software Defined Networking and its Security Threats.Humanitarian Technology Conference (KHI-HTC)(2024), PP 5, Apr2024.

[8] Fotse, Y. S. N., Tchendji, V. K., & Velempini, M. (2024). Federated Learning Based DDoS Attacks Detection in Large Scale Software-Defined Network. IEEE Transactions on Computers (2024), Vol 14, PP – 14, Oct 2024.

[9] Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W. (2023). A survey on federated learning: challenges and applications. International Journal of Machine Learning and Cybernetics (2023), vol 14, PP 513-535, Nov 2023.

[10] Zhu, H., Zhang, H., & Jin, Y. (2021). From federated learning to federated neural architecture search: a survey. Complex & Intelligent Systems (2021), vol 7, PP 639- 657, May 2022.

[11] S.T.Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (ddos) flooding attacks," IEEE communications surveys & tutorials, vol. 15, no. 4, pp. 2046–2069, 2013.

[12] M. Aslan and A. Matrawy, "A clustering-based consistency adaptation strategy for distributed sdn con trollers," in 2018 4th IEEE Conference on Network Softwarization and Workshops (netsoft). IEEE (2018), pp. 441–448

[13] Y. Zhang, L. Cui, W. Wang, and Y. Zhang, "A survey on software defined networking with multiple controllers," Journal of Network and Computer Applications (2018), vol. 103, pp 101–118, 2018.

[14] A. Dixit, F. Hao, S. Mukherjee, T. Lakshman, and R. Kompella, "Towards an elastic distributed sdn controller," ACM SIGCOMM computer communication review (2013), vol. 43, no. 4, pp. 7–12, 2013.