CNN Driven DDoS Detection

A Drill Down Analysis Approach

Venkata Mani Sai Lokesh Kollapudi[1], Ramakrishna Konchada[2], Khuresh Kundurthi[3], Jeevan Babu Maddala[4]

[1]Dept. Computer Science, Vasireddy Venkatadri Institute of Technology, India, Email: kvmslokesh@gmail.com
[2]Dept. Computer Science, Vasireddy Venkatadri Institute of Technology, India, Email: ramakrishnakonchada4@gmail.com
[3]Dept. Computer Science, Vasireddy Venkatadri Institute of Technology, India, Email: kundurthikhuresh64@gmail.com
[4]Asst. Prof., Dept. Computer Science, Vasireddy Venkatadri Institute of Technology, India, Email: jeevan.projects@gmail.com

***Abstract*— With the growing threat of Distributed Denial-of-Service (DDoS) attacks, detecting and mitigating these attacks in real-time is essential for maintaining network security. In this paper, we propose a Convolutional Neural Network (CNN)-based approach to identify and classify DDoS attacks by analyzing network traffic patterns, including the corresponding IP addresses, to differentiate normal traffic from potentially harmful traffic. The model is capable of detecting various attack types, such as TCP SYN, UDP Flood, and ICMP, and classifying the traffic accordingly. Experimental results show that the CNN model achieves high accuracy in classification, outperforming traditional machine learning methods in terms of both training speed and detection efficiency. The model's precision in identifying attack traffic and accurately associating it with the corresponding IP addresses further demonstrates its potential for real-time, large-scale deployment in network security systems.**

Keywords: Attack Identification, Traffic Analysis, Network Protection, Artificial Intelligence, Convolutional Neural Networks (CNN), Distributed Denial of Service (DDoS) Detection.

I. Introduction

A. Background

Distributed Denial of Service (DDoS) attacks continue to pose a significant threat to network security, disrupting services and causing financial losses globally. These attacks exploit the increasing volume and variety of internet traffic, overwhelming target systems with malicious traffic while mimicking legitimate patterns to evade traditional detection methods. As networks evolve with advancements in the Internet of Things (IoT) and cloud computing, the sophistication and scale of DDoS attacks have grown, necessitating robust and adaptive detection mechanisms.

B. Motivation

Conventional DDoS detection methods, such as rule-based systems or statistical analysis, often fail to adapt to evolving attack patterns and require extensive manual tuning. Machine learning (ML) approaches, particularly deep learning, have emerged as effective tools for automatic detection and classification of malicious traffic. Among these, Convolutional Neural Networks (CNNs) are particularly promising due to their ability to process structured data like traffic images, enabling the identification of intricate patterns in network traffic.

This paper presents a system that utilizes CNNs to transform network traffic into 2D traffic images and classify traffic as malicious or benign. Moreover, the system categorizes attack types based on protocols, such as TCP SYN attacks, UDP floods, and ICMP attacks, making it a comprehensive solution for traffic analysis.

To enhance detection precision, a drill-down approach is employed. By focusing only on detected subnets, it minimizes memory consumption and reduces computational overhead, making it suitable for high-volume environments. This iterative refinement allows for precise identification of the destination IP address associated with the attack while avoiding the need to monitor the entire address space continuously.

C. Scope and Contributions

The primary focus of this work is the development of a system that processes traffic data through a specified endpoint and delivers high-accuracy results. Key contributions include:

- A robust methodology to transform raw network traffic into traffic images suitable for CNN-based analysis.
- A CNN architecture capable of achieving 97% accuracy in detecting malicious traffic.
- Protocol-based classification to distinguish attack types based on underlying protocols (e.g., TCP, UDP, ICMP).
- A scalable system designed to handle input traffic data, classify it, and return results with associated IPs and attack types.

This work demonstrates the practical utility of CNNs in DDoS detection and highlights their potential in enabling protocol-based traffic classification for enhanced cybersecurity.

II. Related Work

A. Traditional Methods

Traditional methods for detecting Distributed Denial of Service (DDoS) attacks primarily relied on statistical analysis and rule-based systems. These approaches, such as threshold-based detection and flow-based monitoring, aim to identify anomalous traffic patterns by analyzing packet rates, source IPs, and protocol distributions. For example, coarse-grained monitoring techniques focus on differentiating large flows from legitimate traffic using dynamic thresholds.

While effective for detecting high-volume attacks, these methods struggle with stealthier attacks and dynamic traffic patterns. Furthermore, their dependence on fixed rules makes them less adaptable to evolving attack strategies and diverse network environments.

B. *Machine Learning Approaches*

Machine learning (ML) has emerged as a powerful tool for enhancing DDoS detection, providing automated feature extraction and the flexibility to adapt to emerging threats. Early ML- based approaches, such as those leveraging Support Vector Machines (SVMs), demonstrated high detection accuracy.

"An Evolutionary SVM Model for DDoS Attack Detection in Software Defined Networks" highlights the potential of optimized SVMs for scalable attack detection in SDNs. Deep learning (DL) has further advanced the field, with Convolutional Neural Networks (CNNs) being particularly effective for traffic image analysis.

"HollywooDDoS: Detecting Volumetric Attacks in Moving Images of Network Traffic" employs CNNs to analyze sequential traffic images, achieving high detection accuracy for volumetric attacks. However, this approach focuses on binary classification and does not provide finer granularity regarding attack types or targeted IPs.

Similarly, *"Feasibility Evaluation of Compact Flow Features for Real-Time DDoS Attacks Classifications"* explores lightweight features for real-time detection but relies heavily on handcrafted feature engineering, limiting its adaptability.

C. *Need for Protocol Classification*

While traditional and ML-based approaches have improved DDoS detection, they often lack the ability to provide detailed insights into attack characteristics, such as the specific protocol involved or the exact destination being targeted. Techniques like *"Distributed Denial of Service (DDoS) Attack Detection Using Classification Algorithm"* demonstrate high accuracy in differentiating benign and malicious traffic but do not address the classification of attack types. Furthermore, many methods focus on entire traffic volumes rather than using a refined drill-down approach to isolate malicious traffic efficiently.

This gap underscores the need for systems capable of protocol-based classification, which can identify attack types (e.g., TCP SYN, UDP Flood) and provide detailed analysis of targeted IPs. Such capabilities are crucial for enabling fine-grained traffic analysis and effective mitigation strategies in real-world applications.

III. Proposed Methodology

The proposed detection framework integrates traffic image transformation with convolutional neural networks (CNNs) to efficiently identify and classify Distributed Denial of Service (DDoS) attacks. By leveraging the SDN DDoS dataset, the framework employs iterative granularity refinement, allowing for precise identification of malicious IP addresses. The methodology incorporates a drill-down analysis, which enhances the detection process by providing a deeper examination of suspicious traffic patterns.

Fig. 2 illustrates the workflow of this methodology, showcasing a step-by-step approach that optimizes CNN-based classification for DDoS detection, ultimately improving the accuracy and speed of attack identification.

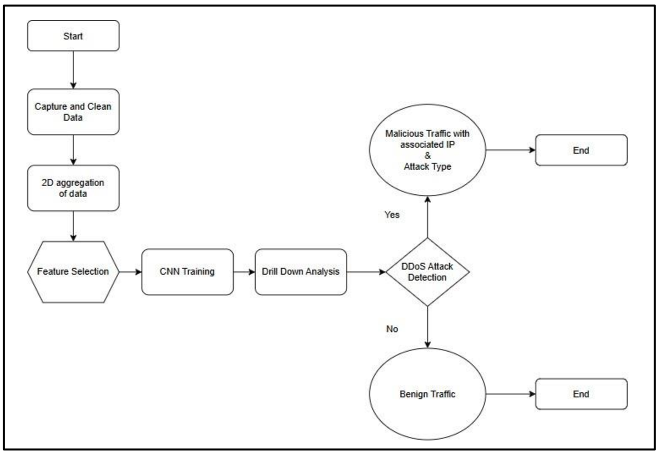

Fig. 1 Workflow

A. *Dataset*

The SDN DDoS dataset was generated using Mininet, with a single Ryu controller managing ten network topologies. It includes both benign and malicious traffic, with malicious traffic further categorized into TCP SYN attacks, UDP Floods, and ICMP attacks.

The dataset contains 23 features, categorized as follows:

- **Extracted Features**: Packet count, byte count, source/destination IPs, port numbers, transferred/received bytes (tx_bytes, rx_bytes), and data transfer rates (tx_kbps, rx_kbps).
- **Calculated Features**: Packet rate, flow entries, flow byte count and packets per flow.

The dataset comprises 104,345 entries gathered within a span of 15000 seconds. Each record is labeled as benign (0) or malicious (1), and additional data can be generated by rerunning the simulation.

B. *Traffic Image Transformation*

The proposed system processes raw network traffic by converting it into structured 2D traffic images. This conversion enables effective feature extraction and classification by utilizing the spatial correlations within the traffic data.

The y-axis of the traffic image represents the complete source IP address space, divided into 256 subnets. The x-axis represents the observed destination IP address range, segmented into smaller subnets. This setup ensures the traffic data is organized in a structured grid for analysis.

Traffic images are generated with a resolution of 256 units, which allows IPv4 destination addresses to be resolved within four iterations of the drill-down process. Each packet is mapped to a pixel within the image based on its source and destination IPs, with the pixel value incremented to represent traffic intensity. Packets outside the monitored IP range are excluded to optimize computational resources and focus the analysis on relevant data.

This transformation produces a compact and visual representation of traffic patterns, enabling the detection of both normal and abnormal behavior using deep learning techniques.

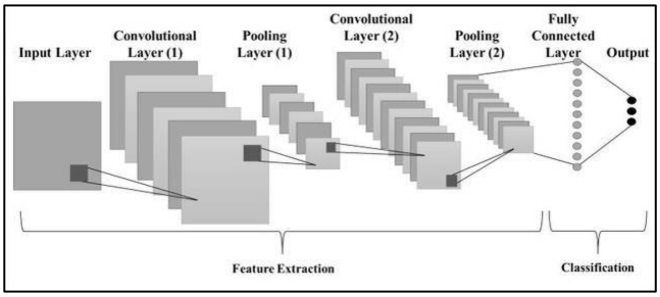

Fig. 2: DDoS Detection Model Architecture

C. *Classification using CNN*

The traffic images are classified through the use of a Convolutional Neural Network (CNN). This network is designed to identify whether traffic is benign or malicious and classify the type of attack based on the protocol used, such as TCP, UDP, or ICMP.

The CNN architecture is composed of several convolutional layers that capture spatial features from the input traffic images. These layers are followed by pooling layers that decrease the dimensionality of the feature maps, preserving essential information while reducing computational load. Fully connected layers then process the extracted features and conduct high-level analysis to produce final classification.

The output layer of the CNN delivers a binary classification for benign or malicious traffic and determines the attack type if malicious traffic is detected. The model is trained using the Adam optimizer and a categorical cross-entropy loss function to facilitate efficient learning and ensure precise outcomes. By automating feature extraction, the CNN effectively handles complex traffic patterns and scales to high-volume datasets.

D. *Drill Down Analysis Approach*

The drill down analysis is a continuous refinement process aimed at progressively narrowing the focus of the monitored traffic from broader patterns to specific subnets and ultimately to individual IP addresses.

In the initial iteration, the system monitors the entire destination IP address space, aggregating traffic into a single image for classification. Once the CNN identifies malicious traffic, the system adjusts the monitoring range and granularity to focus on the identified subnet. This refinement divides the detected subnet into equal segments for further analysis.

With each iteration, the monitored subnet becomes more granular, progressively zooming in on the target until the exact IP address is pinpointed. Typically, IPv4 traffic is resolved within four layers of this process. This iterative approach ensures that computational resources are allocated only to relevant traffic, reducing overhead.

To enhance efficiency and manage dynamic traffic conditions, a Long Short-Term Memory (LSTM) module is incorporated into the system. The LSTM temporarily stores the state of the drill-down process, allowing the system to resume refinement without restarting the analysis if traffic patterns shift or false positives occur. This eliminates the need for a full reset, significantly enhancing processing speed and adaptability.

By combining CNN-based classification with drill-down refinement, the system achieves precise detection while remaining scalable for high-volume traffic environments.

IV. Evaluation and Performance Analysis

A. *Experimental Setup*

The proposed system was evaluated using the SDN DDoS dataset, which encompasses a diverse set of features representing both benign and malicious traffic patterns. The dataset was preprocessed by converting non-numeric values to numeric representations, normalizing the features for consistent scaling, and one-hot encoding the labels for multi-class classification. For model evaluation, the dataset was split into an 80% training set and a 20% testing set to ensure a balanced assessment of the model's performance.

The Convolutional Neural Network (CNN) architecture utilizes convolutional and pooling layers to efficiently extract features. It also includes dropout layers to prevent overfitting. Finally, fully connected layers are used to perform the final classification. The system receives traffic data through a designated endpoint, processes it into a CNN-compatible format, and returns the classification results, classifying the traffic as either benign or malicious and identifying the type of attack, such as TCP SYN, UDP Flood, or ICMP.

B. *Results and Analysis*

The CNN model demonstrated strong performance across both training and validation datasets. Throughout the training, the accuracy of both the training and validation sets steadily increased, while the loss steadily decreased. By the final epoch, the model achieved high accuracy and low loss, indicating its ability to effectively learn from the data and generalize well. The validation accuracy also improved, suggesting that the model was not overfitting and was able to maintain its performance on unseen data.

Throughout the training, the accuracy of both the training and validation sets steadily increased, while the loss steadily decreased, such as Decision Trees or Support Vector Machines (SVM). This efficiency was primarily due to the CNN’s ability to automatically learn features from the raw traffic data, reducing the need for extensive feature engineering and thus minimizing computational time for both training and inference.

Regarding classification, the model achieved strong performance with very few misclassifications. The confusion matrix showed that the model was highly accurate in distinguishing between benign and malicious traffic, confirming its effectiveness in real-time DDoS detection.

The model's performance was assessed using the following key metrics.

Accuracy: The accuracy of the model refers to the percentage of correctly classified instances out of the total instances.

The accuracy is calculated by dividing the sum of True Positives and True Negatives by the total number of instances, which includes the sum of True Positives, True Negatives, False Positives, and False Negatives.

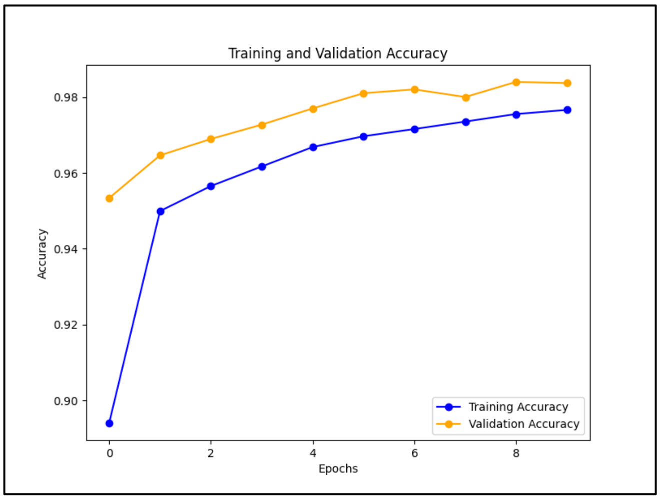

Fig. 3: Model Accuracy Plot

Loss: The loss indicates how closely the model's predictions align with the actual labels. It is computed during training and validation and is minimized over time. The loss is typically calculated using categorical cross entropy function. y_i is the true label $\sim y_i$ is the predicted probability of the class.

$$\text{Loss} = - \sum y_i \log (\sim y_i)$$

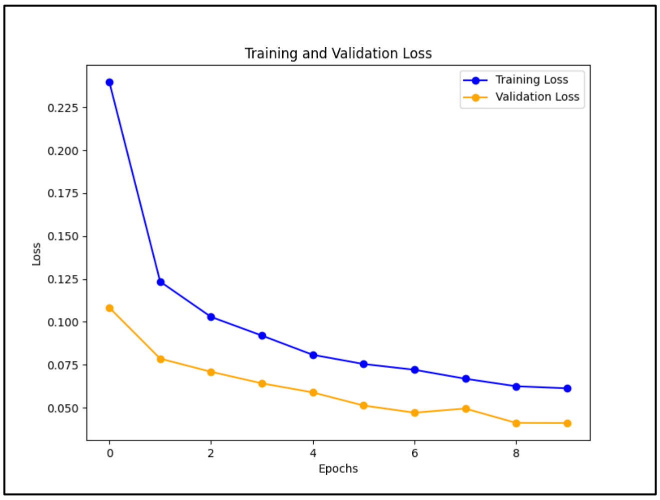

Fig. 4: Model Loss Plot

Validation Accuracy measures how well the model performs on unseen data and *Validation Loss* quantifies the error between the true and predicted values on the validation set.

These metrics confirmed that the model achieved high accuracy and low loss on both training and validation phases, demonstrating its effectiveness for real time DDoS detection.

C. Performance Evaluation and Insights.

As shown in the graphs of Accuracy vs. Epochs and Loss vs. Epochs, the CNN model exhibited consistent improvements in both training and validation accuracy, accompanied by a steady decrease in loss, indicating effective learning and optimization. The increasing validation accuracy further underscores the model's ability to generalize well to unseen data.

The Confusion Matrix clearly illustrates the model's high precision in distinguishing between benign and malicious traffic, with only a minimal number of misclassifications. These results highlight the model's robustness in accurately detecting DDoS attacks.

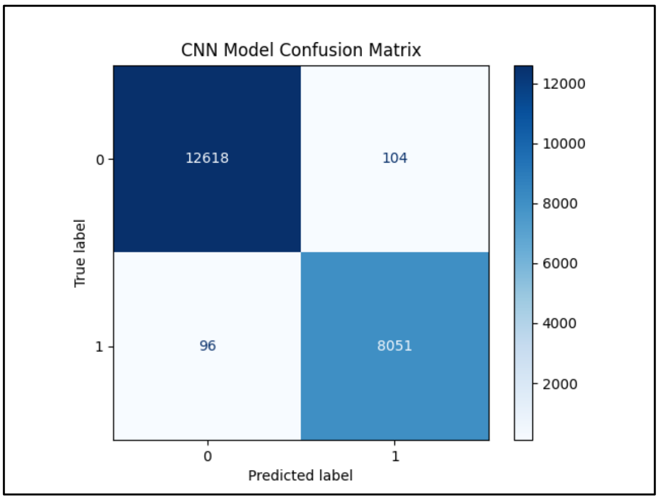

Fig. 5: Confusion Matrix

In terms of time efficiency, the CNN model demonstrated superior performance over traditional machine learning methods like Decision Trees and SVM. By automatically extracting features from raw data, it eliminated the need for feature engineering, thus reducing both training time and inference speed, making it ideal for real-time DDoS detection.

Overall, the CNN model not only delivered high accuracy but also proved more time-efficient than traditional methods, positioning it as a scalable solution for real-time cybersecurity applications.

V. FUTURE WORK

Although the CNN model has demonstrated effective performance in detecting DDoS attacks, there are several areas that could be further explored to enhance its capabilities. A potential avenue for improvement is investigating more advanced architectures, like Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks, which are capable of capturing the temporal patterns in network traffic. This could be particularly beneficial for detecting sophisticated, time- dependent attacks that might not be fully addressed by convolutional networks. Additionally, the training dataset could be expanded by incorporating a broader range of attack types, as well as real-time traffic data, To enhance the model's resilience and its ability to perform well in various types of network environments.

To further optimize the system, the model could be fine-tuned for real-time detection through techniques like model pruning, quantization, and hardware acceleration using GPUs or TPUs. These techniques could significantly reduce inference time, making the model more suitable for high-speed, large-scale environments. Moreover, exploring distributed deployment for DDoS detection, where multiple nodes across the network collaborate to identify attacks, could help mitigate bottlenecks and improve detection accuracy by leveraging distributed computational resources.

Incorporating multiple ML techniques could also help enhance the model's performance. A hybrid approach, combining CNNs with techniques such as ensemble learning or Random Forests, may help enhance classification accuracy and provide more reliable results across a wider range of attack types. Furthermore, it is essential to address the model's resilience to adversarial attacks, where future work could involve adversarial training to ensure the model's robustness against malicious attempts to deceive or evade detection.

VI. Conclusion

This study presents the Drill-Down DDoS Destination Detection approach using a Convolutional Neural Network (CNN) for detecting and classifying Distributed Denial-of-Service (DDoS) attacks. The model analyzes network traffic patterns, including associated IP addresses, to effectively distinguish between benign and malicious traffic and identify various attack types, such as TCP SYN, UDP Flood, and ICMP.

The proposed approach achieved high accuracy, significantly outperforming traditional machine learning methods in both classification precision and computational efficiency. Experimental results demonstrated that the CNN model consistently delivered accurate traffic classification, showcasing its ability to generalize well from both synthetic and authentic traffic data. The model's impressive accuracy and fast evaluation times make it a highly effective tool for detecting DDoS attacks in network security systems.

With its robust performance and scalability, this CNN-based approach offers great potential for real-world applications, providing reliable and efficient detection of DDoS attacks for enhanced cybersecurity.

VII. References

[1] S. Kopmann, T. Krack and M. Zitterbart, "7D: Demonstrating Drill-Down DDoS Destination Detection," 2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN), Stockholm, Sweden, 2024, pp. 1-2, doi: 10.1109/ICMLCN59089.2024.10624795.

[2] H. Liu, Y. Sun, V. C. Valgenti and M. S. Kim, "TrustGuard: A flow-level reputation-based DDoS defense system," 2011 IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, USA, 2011, pp. 287-291, doi: 10.1109/CCNC.2011.5766474.

[3] Silva, M., Marques, J., Gaspary, L., & Granville, L. (2020). Identifying elephant flows using dynamic thresholds in programmable IXP networks. *Journal of Internet Services and Applications*. https://doi.org/10.1186/s13174-020-00131-6A.

[4] A. Alashhab et al., "Enhancing DDoS Attack Detection and Mitigation in SDN Using an Ensemble Online Machine Learning Model," in IEEE Access, vol. 12, pp. 51630-51649, 2024, doi: 10.1109/ACCESS.2024.3384398.

[5] F. Reza, "DDoS-Net: Classifying DDoS Attacks in Wireless Sensor Networks with Hybrid Deep Learning," 2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), Dhaka, Bangladesh, 2024, pp. 487-492, doi: 10.1109/ICEEICT62016.2024.10534545.

[6] B. Nagpal, P. Sharma, N. Chauhan and A. Panesar, "DDoS tools: Classification, analysis and comparison," 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2015, pp. 342-346.

[7] M. A. Saleh and A. Abdul Manaf, "Optimal specifications for a protective framework against HTTP-based DoS and DDoS attacks," 2014 International Symposium on Biometrics and Security Technologies (ISBAST), Kuala Lumpur, Malaysia, 2014, pp. 263-267, doi: 10.1109/ISBAST.2014.7013132.

[8] Kopmann, S., Heseding, H., & Zitterbart, M. (2022). HollywooDDoS: Detecting Volumetric Attacks in Moving Images of Network Traffic. *2022 IEEE 47th Conference on Local Computer Networks (LCN)*, 90-97. https://doi.org/10.1109/LCN53696.2022.9843465.

[9] M. H. Rohit, S. M. Fahim and A. H. A. Khan, "Mitigating and Detecting DDoS attack on IoT Environment," 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON), Dhaka, Bangladesh, 2019, pp. 5-8, doi: 10.1109/RAAICON48939.2019.5.

[10] B. Jia and Y. Liang, "Anti-D chain: A lightweight DDoS attack detection scheme based on heterogeneous ensemble learning in blockchain," in China Communications, vol. 17, no. 9, pp. 11-24, Sept. 2020, doi: 10.23919/JCC.2020.09.002.

[11] M. Hassan, K. Metwally and M. A. Elshafey, "ZF-DDOS: An Enhanced Statistical-Based DDoS Detection Approach using Integrated Z-Score and Fast-Entropy Measures," 2024 6th International Conference on Computing and Informatics (ICCI), New Cairo - Cairo, Egypt, 2024, pp. 145-152, doi: 10.1109/ICCI61671.2024.10485097.

[12] Fei Wang, Xiaofeng Hu, Xiaofeng Wang, Jinshu Su and Xicheng Lu, "Unfair rate limiting on traffic aggregates for DDoS attacks mitigation," IET International Conference on Information Science and Control Engineering 2012 (ICISCE 2012), Shenzhen, 2012, pp. 1-5, doi: 10.1049/cp.2012.2448.

[13] W. H. A. Muragaa, "A hybrid scheme for detecting and preventing single packet Low-rate DDoS and flooding DDoS attacks in SDN," 2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA), Benghazi, Libya, 2023, pp. 707-712, doi: 10.1109/MI-STA57575.2023.10169712.

[14] P. Prathap and S. Duttagupta, "AI-Enabled Fast Detection of DDoS and Adversary DDoS Attacks in SDN," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-7, doi: 10.1109/ICCCNT56998.2023.10306608.

[15] A. Ramzy Shaaban, E. Abdelwaness and M. Hussein, "TCP and HTTP Flood DDOS Attack Analysis and Detection for space ground Network," 2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES), Cairo, Egypt, 2019, pp. 1-6, doi: 10.1109/ICVES.2019.8906302.

[16] J. A. Pérez-Díaz, I. A. Valdovinos, K. -K. R. Choo and D. Zhu, "A Flexible SDN-Based Architecture for Identifying and Mitigating Low-Rate DDoS Attacks Using Machine Learning," in IEEE Access, vol. 8, pp. 155859-155872, 2020, doi: 10.1109/ACCESS.2020.3019330.

[17] J. Bhayo, S. Hameed and S. A. Shah, "An Efficient Counter-Based DDoS Attack Detection Framework Leveraging Software Defined IoT (SD-IoT)," in IEEE Access, vol. 8, pp. 221612-221631, 2020, doi: 10.1109/ACCESS.2020.3043082.

[18] N. M. Yungaicela-Naula, C. Vargas-Rosales, J. A. Perez-Diaz, E. Jacob and C. Martinez-Cagnazzo, "Physical Assessment of an SDN-Based Security Framework for DDoS Attack Mitigation: Introducing the SDN-SlowRate-DDoS Dataset," in IEEE Access, vol. 11, pp. 46820-46831, 2023, doi: 10.1109/ACCESS.2023.3274577.

[19] S. Dong, K. Abbas and R. Jain, "A Survey on Distributed Denial of Service (DDoS) Attacks in SDN and Cloud Computing Environments," in IEEE Access, vol. 7, pp. 80813-80828, 2019, doi: 10.1109/ACCESS.2019.2922196.

[20] Introducing the LATAM-DDoS-IoT Dataset," in IEEE Access, vol. 10, pp. 106909-106920, 2022, doi: 10.1109/ACCESS.2022.3211513.