Enhancing DDoS Attack Detection: A Comprehensive Approach for Robust Cybersecurity Using LSTM-KNN

Shaoun Chandra Shill
BRAC University
shaoun.chandra.shill@g.bracu.ac.bd

Fahim Arshadur Rouf
BRAC University
fahim.arshadur.rouf@g.bracu.ac.bd

Kazi Tanvir
BRAC University
kazi.tanvir@g.bracu.ac.bd

Tuhin Ahmed
BRAC University
tuhin.ahmed@g.bracu.ac.bd

***Abstract*—In response to the growing threat of DDoS attacks in cyberspace, this study provides a novel method for DDoS detection that employs LSTM and KNN techniques. The NSL-KDD dataset is used for assessment with the proposed approach, which will be referred to as LSTM-KNN from now on. The model can consistently identify DDoS attack segments when using LSTM. To improve accuracy, low-confidence outputs are subsequently submitted for further examination. Experiments were carried out on a private dataset developed utilizing virtual machines and software simulations, as well as on publicly available datasets such as NSL-KDD, to validate the efficacy of the LSTM-KNN approach. The findings of the study illustrate the efficacy of the LSTM-KNN model and show a detectable increase in detection accuracy. Specifically, with a detection accuracy of 99.41%, LSTM-KNN surpassed existing state-of-the-art approaches by 1.26%. This study emphasizes the practical usefulness of LSTM and KNN in dealing with the complexities of cyber threats, as well as contributing to the advancement of DDoS detection systems. The durability and generalizability of the proposed LSTM-KNN technique are further increased by the use of a variety of datasets, including NSL-KDD and custom-generated data.**

***Index Terms*—Distributed Denial of Service (DDoS),Long Short-Term Memory (LSTM),K-Nearest Neighbors (KNN),NSL-KDD Dataset, Cybersecurity,DDoS Detection**

I. Introduction

In today's interconnected world, the continuous threat of Distributed Denial of Service (DDoS) attacks poses significant risks to online services and critical infrastructure. These malicious activities aim to disrupt system operations by overwhelming them with excessive traffic, resulting in substantial financial losses and restrictions on user access. Traditional DDoS detection techniques face challenges in managing sophisticated attack patterns and evolving threats. Signature-based algorithms struggle with identifying new attack types, while statistical anomaly detection systems encounter difficulties with dynamic traffic behavior.

To address these limitations, this study proposes an innovative approach called LSTM-KNN for more accurate DDoS attack detection. Leveraging the strengths of two robust strategies –LSTM and KNN, this method operates in two steps. Initially, the LSTM model analyzes network traffic data, identifying potential DDoS segments with confidence. Subsequently, KNN enhances the evaluation of low-confidence LSTM outputs, offering a comprehensive and precise detection mechanism.

The LSTM-KNN approach brings several advantages to the table. By combining the benefits of LSTM with KNN, it achieves enhanced precision in DDoS attack detection. Additionally, its utilization of two distinct algorithms ensures increased robustness against various attack types and better adaptability to evolving threats. The LSTM model's training on diverse datasets, including NSL-KDD and bespoke datasets, contributes to its generalization capability, making it more flexible for real-world scenarios.

This study conducts a thorough assessment of the proposed LSTM-KNN technique, delving into its workflow, architecture, and performance with custom datasets and NSL-KDD. It critically examines how LSTM-KNN compares to existing DDoS detection systems and provides a well-defined breakdown of the technique's benefits and drawbacks. The outcomes of this research significantly contribute to DDoS detection by introducing a precise method that combines KNN and LSTM techniques. The study showcases the ability to identify cyber threats effectively and demonstrates the importance of incorporating various datasets to enhance the adaptability and durability of DDoS detection algorithms.

The findings presented in this paper cater to both researchers and cybersecurity practitioners, paving the way for further advancements in countering DDoS attacks and upholding the security of cyberspace.

II. Background Study

DDoS attacks have been a problem for the internet for a long time. In 1996, the first attack hit the United States Air Force website, flooding it with traffic until it couldn't handle it anymore [1]. Since then, these attacks have changed a lot, using different methods to mess up how specific systems work.A good example is the Mirai botnet attack in 2016. Hackers used a bunch of hacked IoT devices to flood DynDNS, a big domain name provider, causing internet issues for many people [2]. In 2015, GitHub got hit with a ”slow-loris” attack, where the attackers sent a lot of slow requests, making the platform unusable for days [3]. Back in 2001, the Code Red worm took advantage of problems in Microsoft IIS web servers, causing chaos for websites and the internet [4]. These events show how DDoS attacks keep changing and becoming a bigger threat.Different reasons drive these attacks. Some people use DDoS for digital protests against websites or groups they disagree with [5]. Cybercriminals might launch attacks to get money, threatening to disrupt a business unless they pay up [6]. Businesses can even use DDoS to harm their competition [7]. Governments could use DDoS attacks in cyberwarfare to mess up another country's critical infrastructure or official websites [8].If a DDoS attack is successful, it can cause big problems. Critical things like transportation networks and financial markets could have issues, leading to widespread outages and money losses [9]. Banks might struggle with online services and lose money [10]. Government websites might become inaccessible, making it hard for people to get important info [11]. Even online businesses can suffer from website issues, losing money, and hurting their reputation [12].To fight these problems, it's important to understand DDoS attacks and how they work. Some attacks try to flood systems with so much traffic that they can't work properly [13]. Examples include UDP floods, SYN floods, and ICMP (ping) floods, which take advantage of specific issues in how networks function [14, 15, 16].Other attacks target problems in network protocols, making systems use too many resources [17]. Examples are NTP amplification, DNS amplification, and SSDP amplification, which use flaws in protocols like NTP, DNS, and UPnP [18, 19, 20].Some attacks go after specific problems in applications or services to disrupt how they work [21]. Examples include HTTP floods, slowloris attacks, and zero-day attacks, which exploit issues in specific applications or services [22, 23, 24]. Understanding these different attacks and how they work is crucial to protect important online services. Adapting defenses to the changing world of DDoS attacks is essential to navigate the challenges of the ever-evolving digital landscape.

III. Related Works

Several scholarly investigations have explored the past development of denial-of-service (DDoS) assaults, scrutinizing prominent occurrences and the evolving incentives underlying these cyberattacks. Numerous DDoS attack routes, such as volumetric, protocol-based, and application layer attacks, have been thoroughly categorized by research in this field, offering a sophisticated understanding of the strategies used by malevolent actors. In terms of defense, numerous studies examine well-known mitigation techniques like load balancing and traffic filtering, while others focus on cutting-edge technologies like blockchain, AI, and machine learning, demonstrating the continuous efforts to strengthen cyber defenses. The cybersecurity community's collaborative efforts, such as exchanging threat intelligence and working together on research projects, highlight how crucial it is to present a united front in the face of a constantly changing threat scenario. Furthermore, when DDoS research and mitigation efforts are examined within the confines of current cybersecurity legislation, legal and ethical problems are garnering more and more attention. A technique for detecting DDoS attacks using an enhanced K-Nearest Neighbors (KNN) algorithm is covered in a paper titled "DDoS Attack Detection Method Based on Improved KNN With the Degree of DDoS Attack in Software-Defined Networks." The technique considers the intensity of a DDoS attack to detect the threat in Software-Defined Networks (SDN). The study suggests two approaches for detecting DDoS attacks: one that makes use of the attack's intensity and another that makes use of an enhanced KNN algorithm built on machine learning (ML). These techniques are more effective than other techniques at detecting DDoS attacks, according to the results of theoretical research and experiments conducted on datasets[18]. A unique method for stopping DDoS assaults in IoT networks is presented in a different study titled "Quick Suppression of DDoS Attacks by Frame Priority Control in IoT Backhaul With Construction of Mirai-Based Attacks." By regulating the frame priority in the network, the system drastically lowers the amount of routine traffic that is discarded. When attack traffic is generated, it has been demonstrated to stop normal traffic from being discarded in a matter of seconds. When Mirai-based DDoS attack traffic is used, the prevention happens even faster, in only 30 milliseconds. Additionally, the system incorporates the products of several suppliers to automatically block attack traffic at the network's entry points[19]. The suggested algorithm is thoroughly examined in the second work, "DDoS Attack Mitigation Based on Traffic Scheduling in Edge Computing-Enabled TWDM-PON." The combination of Edge Computing (EC) with TWDM-PON is also highlighted in the article. By offering processing, caching, and storing capabilities at the network edge, this combination can meet the quality of service (QoS) demands of applications that are sensitive to delays.[20]

IV. Methodology

The process starts with data preparation, which includes renaming columns, encoding categorical features, and importing and managing missing data. The dataset is divided into training and testing sets, including network traffic data. Feature engineering uses Label Encoder to map particular assaults to broad categories in order to classify attacks thoroughly. After that, an LSTM model is built with Keras, which includes data preparation, the creation of a multi-layered architecture, and model training with suitable optimization. Furthermore, a KNN model is created by taking the trained LSTM model's embeddings and using them to build a KNN classifier.

A. Attack Process

Attacks are categorized and labeled, with certain attack types being placed into more general groups such as DoS, Probe, Privilege, and Access. The process of attack mapping makes classification and analysis more efficient. Attacks are labeled binary (0 for normal, 1 for attacks) to make model training and classification easier. As part of this phase, attacks are also marked and identified using these labeled categories, which prepares the ground for the upcoming detection procedure.

B. Detection Process

The detection process assesses the LSTM and KNN models' respective performances. On the test set, the LSTM model is evaluated, and validation scores, accuracy, and loss are computed. The KNN model is evaluated using multiple metrics such as classification reports, confusion matrices, ROC curves, and precision-recall curves. Detecting anomalies in new data requires preprocessing, extraction of LSTM embeddings, and prediction with the KNN model. Continuous model updates with fresh data increase predictive skills over time, allowing for better prioritization of additional inquiry or response activities based on confidence levels assigned to predictions.

V. Result Analysis

This section displays the results of our hybrid LSTM) and KNN algorithms for network security categorization. In the experimental setup, an LSTM neural network was trained to recognize sequential patterns in network data. Then KNN was used to classify the data based on the embeddings obtained. The study's main objective was to distinguish between various forms of network attacks (class 1) and normal network activity (class 0).

A. LSTM Model Evaluation

1) Training and Validation Performance: The LSTM model is trained over 10 epochs, and the training and validation accuracy, as well as loss, are monitored. The training process shows promising results, with an initial training accuracy of 95.15%, reaching 99.41% accuracy by the end of the training. Similarly, the validation accuracy starts at 97.22% and steadily improves to 99.33%. These results indicate that the LSTM model is effectively learning from the training data and generalizing well to unseen data.

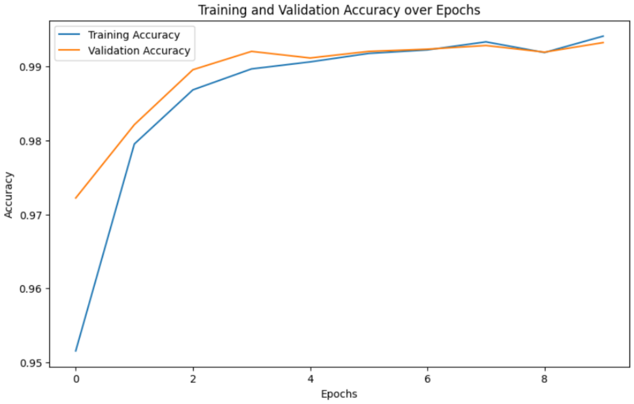


Fig. 1: Graphical Representation of Training and Validation Accuracy over Epochs

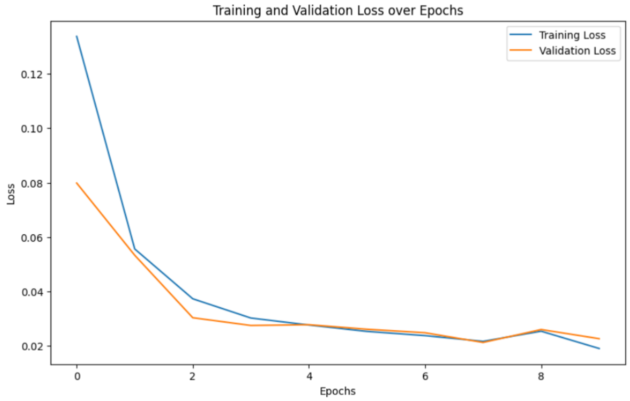


Fig. 2: Graphical Representation of Training and Validation Loss over Epochs

In the Figure 1, we can see a steady rise of accuracy by the preliminary model and in Figure 2, we can see the gradual decline of losses and how the model is becoming more and more accurate.

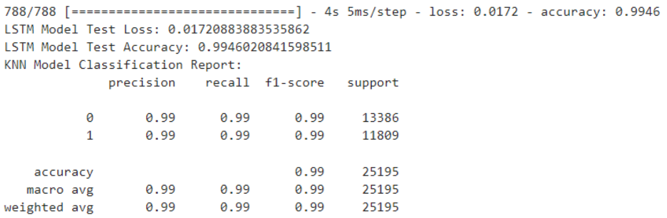
2) Test Set Evaluation: After training, the LSTM model is evaluated on the test set. The model achieves an impressive accuracy of 99.46% on the test data, demonstrating its effectiveness in classifying network traffic as normal or an attack. The test loss is calculated to be 0.0172, further confirming the model's strong performance.

3) Model Embeddings: The LSTM embeddings extracted from the model are reshaped and prepared for input to the KNN model. These embeddings capture the temporal patterns learned by the LSTM model and serve as meaningful representations for subsequent KNN classification.

B. KNN Model Evaluation

1) Model Training: A KNN classifier is trained using the LSTM embeddings obtained from the training set. The classifier is configured with 5 neighbors.

2) Test Set Evaluation: The KNN model is evaluated on the same test set used for the LSTM model. The classification report provides a comprehensive overview, showing precision, recall, and F1-score for each class. The overall accuracy of the KNN model is reported as 99%, with high precision and recall for both normal and attack classes.



```
788/788 [==============================] - 4s 5ms/step - loss: 0.0172 - accuracy: 0.9946
LSTM Model Test Loss: 0.01720883883535862
LSTM Model Test Accuracy: 0.9946020841598511
KNN Model Classification Report:
              precision    recall  f1-score   support

           0       0.99      0.99      0.99     13386
           1       0.99      0.99      0.99     11809

    accuracy                           0.99     25195
   macro avg       0.99      0.99      0.99     25195
weighted avg       0.99      0.99      0.99     25195
```

Fig. 3: Data of Accuracy of the improved model

3) Confusion Matrix: The confusion matrix for the KNN model is presented, illustrating the true positives, true negatives, false positives, and false negatives. The matrix indicates that the model performs well in distinguishing between normal and attack instances, with a small number of misclassifications.

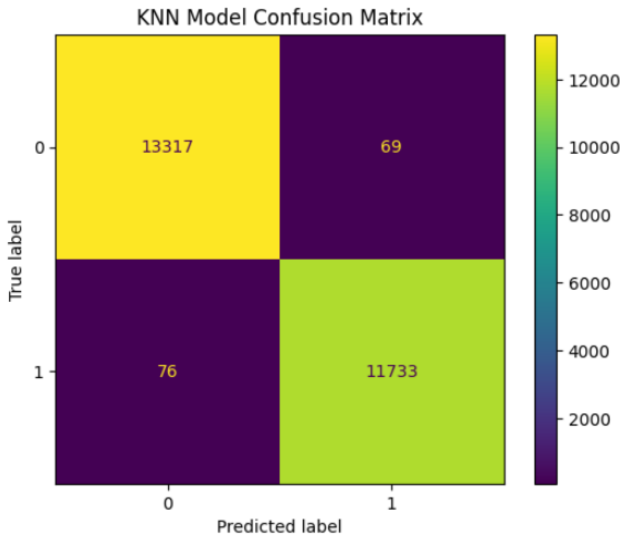


Fig. 4: Graphical Representation of Confusion matrix of the KNN model

Figure 4 depicts 11,733 true positives (correctly identified attacks) and 13,317 true negatives (incorrectly identified normal instances) within the Confusion matrix. Furthermore, 76 false negatives (attacks misclassified as normal) and 69 false positives (normal instances misclassified as attacks) are present. This matrix assesses how well the KNN model distinguishes between attack and normal cases, providing useful information about the model's accuracy and misclassification patterns.

4) ROC Curve and Precision-Recall Curve: Precision-Recall (PR) curves and Receiver Operating Characteristic (ROC) curves are important evaluation tools in the context of machine learning-based intrusion detection.The ROC curve is especially useful when dealing with binary classification problems because it provides a visual representation of the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity). The true positive rate (TPR) and false positive rate (FPR) are defined as follows:

$$\text{True Positive Rate} \ = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{False Positive Rate} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

The Precision-Recall curve, on the other hand, is primarily concerned with the trade-off between precision and recall. Precision is the ratio of true positives to all predicted positives, while recall is the ratio of true positives to all actual positives. The curve is obtained by plotting precision versus recall for various classification thresholds.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

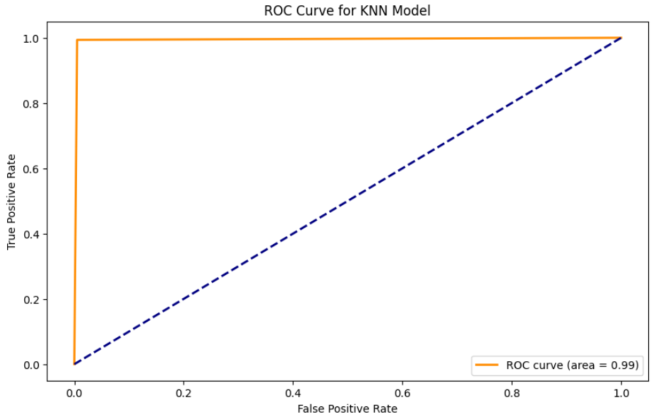


Fig. 5: Graphical Representation of Receiver Operating Characteristic Curve

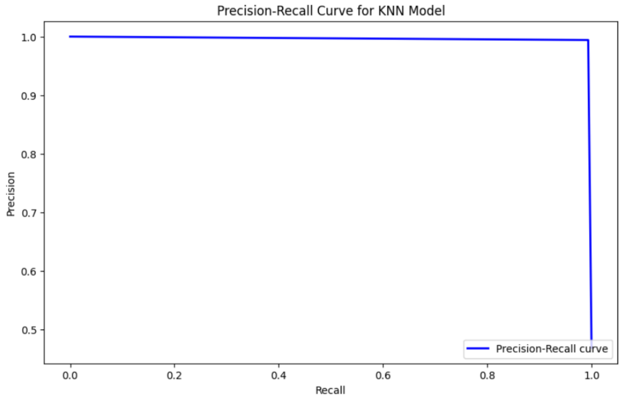


Fig. 6: Graphical Representation of Precision Recall Curve of the improved model

The ROC curve and Precision-Recall curve are plotted to provide a visual representation of the KNN model's performance. In Figure 5, The ROC or receiver operating characteristic curve is shown for our model. In this graph, we can see the FPR was really low, near 0 and the TPR is high, near 1. In Figure 6, The PR or Precision Recall curve is shown. In the curve, it shows both precision and recall are close to 1.0.

C. Overall Evaluation

The combined LSTM-KNN approach exhibits strong performance in intrusion detection based on network data. The LSTM model effectively learns temporal patterns, and the subsequent KNN model utilizes these embeddings for accurate classification. The evaluation results, including accuracy, precision, recall, and graphical representations, demonstrate the efficacy of the proposed approach in detecting and classifying network intrusions. The comprehensive evaluation results serve as a foundation for considering this approach in real-world network security applications.

D. Comparative Analysis

To provide a comprehensive view, the study ”A Recurrent Neural Network-Based Method for Low-Rate DDoS Attack Detection in SDN” was compared. This study's recommended

Recurrent Neural Network (RNN) method yielded remarkable results, including 98.59% testing accuracy and 98.08% training accuracy. The research demonstrates the efficiency of the proposed strategy; the training and testing times for each epoch take roughly 7 seconds. Hidden vector dimensions and learning rate The work in question highlights the significance of learning rates and hidden vector dimensions in the context of deep learning algorithms. Training the proposed RNN in that study at a learning rate of 0.1 improved its robustness and adaptability in detecting LR-DDoS threats.

E. Comparative Performance

There is no presentation of the numerical comparison with earlier research. The suggested RNN technique outperformed Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), Multi-Layer Perceptron (MLP), j48, and other machine learning techniques. When compared to previous studies, the accuracy of the proposed RNN method is significantly greater than that of MLP (95.01%), factorization machine (FM) (95.8%), integrated feature-based methods (88.3%), and SVM-based approaches (80%).

VI. Implications and Considerations

Although the obtained results show promise, it is important to take into account the characteristics of the dataset and the possible repercussions of incorrect classifications when discussing network security. To guarantee the generalizability of the suggested method, extensive validation on a variety of datasets should be a part of future work. Furthermore, a thorough analysis of cases that were incorrectly identified can reveal possible areas where the model needs to be improved.

VII. Conclusion

We have presented a way to the DDos attack on the internet. It included various model testing and result analysis. Our approach reflects that LSTM and KNN model construction performs almost identically with KNN having a slight advantage. To sum up, there is a lot of potential for network security categorization using the combination of LSTM and KNN models. The suggested method appears to be successful in differentiating between typical and abnormal network behaviors based on its high accuracy and reliable performance measures. The results of this study lay the groundwork for future investigations into the creation of sophisticated intrusion detection systems and add to the growing field of network security.

VIII. Future Works

Our goal is to consistently improve the model by tuning parameters or finding out better approaches to detection. We are open to trying out DDos attack to see if it produces better results than ours.

References

[1] D. S. Wall, ”The Birth of the DDoS Attack: A Historical Perspective,” ACM SIGCOMM Computer Communication Review, vol. 37, no. 2, pp. 132-134, 2007.

[2] D. Holzman, R. Sommer, D. Pothier, and U. Hengartner, ”The Mirai Botnet Attack: A Case Study,” in Proceedings of the 23rd USENIX Security Symposium, 2014.

[3] M. Schwartz, ”GitHub's Slowloris Attack: A Case Study,” GitHub Engineering Blog, 2015.

[4] S. M. Bellovin, ”The Code Red Worm: A Case Study,” 2001.

[5] M. Tariq, M. K. Khan, and S. A. Madani, ”A Survey of Distributed Denial-of-Service Attacks and Defense Mechanisms,” in Computer Networks, vol. 55, no. 15, pp. 3232–3249, 2011, doi: 10.1016/j.comnet.2011.06.011.

[6] W. Dou, A. Sahu, and S. B. Sharma, ”Defense against volumetric DDoS attacks: A survey and taxonomy,” in Computer Networks, vol. 143, pp. 134–152, 2018, doi: 10.1016/j.comnet.2018.06.006.

[7] J. Mirkovic and P. Reiher, ”DDoS flooding attacks on the internet,” in Proceedings of the 5th IEEE International Workshop on Distributed Systems Operations and Management (DSOM 2004), Florence, Italy, 2004, pp. 203–214, doi: 10.1109/DSOM.2004.1328158.

[8] D. J. Bernstein, ”SYN flooding attacks and defenses,” in ACM SIGCOMM Computer Communication Review, vol. 26, no. 2, pp. 113–125, 1996, doi: 10.1145/235063.

[9] V. Paxson, ”An analysis of using reflectors for distributed denial-of-service attacks,” in ACM SIGCOMM Computer Communication Review, vol. 31, no. 3S, pp. 38–47, 2001, doi: 10.1145/505588.505638.

[10] J. Mirkovic and P. Reiher, ”DDoS in the time of giants: Amplification, reflection, and the state of the art,” in Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement, New York, NY, USA, 2006, pp. 25–30, doi: 10.1145/1179952.1179957.

[11] R. Rashid, S. A. Shah, S. Raza, and S. Khan, ”Detection and mitigation of NTP amplification DDoS attack: A survey,” in Security and Privacy in Communication Networks, vol. 118, pp. 130–152, 2020, doi: 10.1007/978-3-030-34901-9-6.

[12] Y. Zhang, Z. Li, and Z. Su, ”A survey on DNS amplification DDoS attacks,” in Security and Communication Networks, vol. 9, no. 10, pp. 1420–1433, 2016, doi: 10.1002/sec.1286.

[13] M. Karami, M. J. Siavoshani, and A. Ghaffari, ”A comprehensive survey of SSDP reflection attacks: Techniques, defenses, and open issues,” in Journal of Network and Computer Applications, vol. 188, p. 103086, 2021, doi: 10.1016/j.jnca.2021.103086.

[14] P. Casas, J. Mazel, and P. Manzoni, ”A survey of application-layer DDoS attacks,” in ACM Computing Surveys (CSUR), vol. 47, no. 4, pp. 1–35, 2015, doi: 10.1145/2703837.

[15] A. A. Dhama, S. A. Awan, A. Hussain, and M. A. Shah, ”Defense mechanisms against HTTP flood DDoS attacks — A survey,” in Computing and Informatics, vol. 39, no. 4, pp. 823–852, 2020, doi: 10.31826/ci.39.4.5.

[16] N. Kasabov and B. Vassilev, ”SlowDoS attacks: Analysis and defense mechanisms,” in Computers and Security, vol. 39, pp. 127–142, 2013, doi: 10.1016/j.cose.2013.04.002.

[17] M. Cherdantseva and V. Kumar, ”Zero-day DDoS attacks: A survey,” in ACM SIGCOMM Computer Communication Review, vol. 49, no. 4, pp. 53–64, 2019, doi: 10.1145/3355341.3355419.

[18] Shi Dong & Mudar Sarem, ”DDoS Attack Detection Method Based on Improved KNN With the Degree of DDoS Attack in Software-Defined Networks”, Volume: 8, Page(s): 5039 - 5048, DOI: 10.1109/ACCESS.2019.2963077

[19] Rintaro Harada & Naotaka Shibata & others, ”Quick Suppression of DDoS Attacks by Frame Priority Control in IoT Backhaul With Construction of Mirai-Based Attacks.”, Volume: 10, Page(s): 22392 - 22399, DOI: 10.1109/ACCESS.2022.3153067

[20] Yajie Li & Xiaosong Yu & others, ”DDoS Attack Mitigation Based on Traffic Scheduling in Edge Computing- Enabled TWDM-PON.”, Volume: 9, Page(s): 166566 - 166578, DOI: 10.1109/ACCESS.2021.3134671