International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

A Stratified Deep Detection Pipeline For Ddos Threats Leveraging Recursive Dimensionality Reduction, Probabilistic Latent Encoding, And Temporal Contextualization Networks

Gibi K S[1]*, Dr. S. Nithya[2]
[1*]Research Scholar, Department of Computer Science, Park's College, Tirupur, gibikakkanate297@gmail.com
[2]Assistant Professor, Department of Computer Science, AVP College of Arts & Science, Tirupur, drnithyasundaram23@gmail.com

Abstract: *Intelligent and adaptable detection systems are required due to the increasing complexity of Distributed Denial of Service (DDoS) attacks. This paper discuss an innovative blended approach for detecting DDoS threats which uses an Attention-Enriched Transfer Learning (AETL)[20] framework with Variational Autoencoders (VAEs)[19]. In this model, we aimed to enhance anomaly detection. The model utilizes Deep Packet Inspection (DPI), sophisticated Recursive Feature Elimination (RFE), and temporal pattern recognition by Long-Short-Term Memory (LSTM) networks. To optimize the model's efficiency by differentiating between malicious and benign flows, VAEs are used, hence the unsupervised learning of latent representations of network traffic is done. These new experimental findings utilized the most recent datasets, such as CIC-DDoS2020, CIC-DDoS2019, CIC-DDoS2017 and TON_IoT, show enhanced detection accuracy, F1-score, and response time. This architecture is shown to scale efficiently while adapting to emerging attack vectors.*
Keywords: *Variational Autoencoder (VAE), Transfer Learning, Deep Packet Inspection, LSTM, Feature Elimination*

1.INTRODUCTION

DDoS attacks pose a continual threat to network infrastructure by blocking services with illegitimate traffic. With modern networks characterized by high throughput, encryption, and diverse traffic types, traditional static detection systems are increasingly inadequate. Nowadays the attacks are directed towards IoT systems which leads to serious damage to IoT devices. According to the recent estimate, by 2025, 32 billion IoT devices will be active in which 50% will be unprotected. [1], [2]. Therefore Iot devices are considered as the perfect candidate for the exploitation of such attacks [18] In addition to vulnerabilities in IoT environments, threats span across various layers of the system architecture. At the physical and data link layers, insecure interfaces and weak authentication mechanisms can lead to Sybil and spoofing attacks, compromising node identity and trust. Meanwhile, the network layer is particularly susceptible to advanced threats such as replay, sinkhole, and wormhole attacks, which can disrupt data routing and enable eavesdropping or data manipulation. These multi-layer vulnerabilities not only compromise data integrity and availability but also serve as entry points for attackers to aggregate compromised devices into large-scale botnets. Such botnets are often weaponized to execute devastating Distributed Denial of Service (DDoS) attacks, overwhelming target servers with massive traffic volumes and causing significant disruption to critical online services [3]. Emerging techniques using deep learning and generative models offer significant advantages in generalizing from evolving attack signatures.

Table-1 shows the estimated loss value of major DDoS attacks from 2020-2024.

Year	Attack Case	Symbolic Loss Value*	Estimated Financial Impact (USD)
2020	AWS Attack	10	~$1.5 million – $2.0 million
2020	Miami Schools	2	~$100,000 – $200,000
2022	Ukraine Banks	8	~$1.0 million – $1.5 million
2022	Romania Govt	4	~$300,000 – $500,000
2023	India G20 Campaign	5	~$500,000 – $700,000
2024	Internet Archive	7	~$800,000 – $1.2 million
2024	Cloudflare ISP Attack	12	~$2.5 million – $3.5 million (record)
2024	Italy Govt & Airports	6	~$600,000 – $900,000

* ***Symbolic value 1*** *roughly corresponds to* ***$100,000–$200,000****.*

Table-1: Symbolic and Monetary Loss Estimation for Recent DDoS Incidents

The growing scale, frequency, and variety of targets in DDoS attacks emphasize the critical necessity for sophisticated and flexible detection and prevention strategies. Methods like unsupervised learning, variational

autoencoders, and federated detection frameworks are increasingly vital in addressing emerging threats in real-time [4].

This paper builds on our earlier research, there we integrated LSTM, RFE ,DPI, AETL along with Q-learning for DDoS detection as well as mitigation[21]. In the current study, we integrate a combined framework joining the power of Variational Autoencoders (VAEs). This aims to overcome earlier limitations related to the detection and mitigation. Here we used VAE for latent feature extraction and Attention-Enriched Transfer Learning (AETL) for dynamic prioritization of traffic features. Unlike reinforcement learning-based methods, VAEs enable the unsupervised learning of data distributions, making the detection model robust to novel or zero-day attacks [5]. The unsupervised learning methods can directly process raw network data. The collection of data for all variants of DDoS attacks is very challenging and time-consuming [6]. The tactics of attacks are changed frequently. Also, the unsupervised model can efficiently adapt to large network traffic data [7].

2. METHODOLOGY

2.1 Overview Of Previous Model

In our previous research we developed a hybrid model for the DDoS attack detection. Its done by combining Q-learning, LSTM networks, RFE and deep packet inspection. The main goal was to improve the accuracy in anomaly detection by reducing the irrelevant features, so that we can improve the detection time.

The preprocessing was started with DPI, and then with the RFE, the result was a refined set of data with only selecting the relevant features. Q-learning has a vital role in detecting the suspicious patterns and LSTM is for learning the nature of traffic.

2.2 Enhancements In The Current Model

In this work, we enhance our hybrid DDoS detection model by the main component variational Autoencoder (VAE). This improves the accuracy and ability to detect more complex attacks in any kind of network environments.bby introducing VAE, which brings probabilistic perspective to the feature learning. Q-learning was using rule based decision logic, but VAE is identifying the anomalies based on reconstruction of the compressed representation of data. This helps to detect the unseen attacks even.

Additionally ATEL helps to find out the most relevant features by joining the attention mechanism with transfer learning. This helps to better generalization of different datasets and attack patterns. Together DPI, RFE (enables to extract traffic feature and reduce redundancy) selects the relevant features, and pass them through VAE for encoding and estimation of anomaly. LSTM capture the temporal dependency and then the AETL layer refines the learning process. The combination gives more powerful results and an adaptable detection system with high accuracy and speed.

2.3 Data Preprocessing

2.3.1. Deep Packet Inspection (DPI)

DPI is a method used to inspect the Contents of the data packet which is traversing through the network.. Unlike basic packet inspection (which only checks headers), DPI can deeply check the content of the data, also for the detection of any unexpected moves of the network traffic [8]. This helps to generate the details of the header and the data features :

Header features: Source and Destination IP addresses and port numbers, protocol type, TTL, header length, total length, flags, sequence and acknowledgment numbers (TCP), window size, checksum, and more.

Payload features: byte patterns, payload size, entropy, known content signatures, application-level protocols (e.g., HTTP, DNS), command or keyword patterns, payload hashes, and other indicators of malicious behavior.

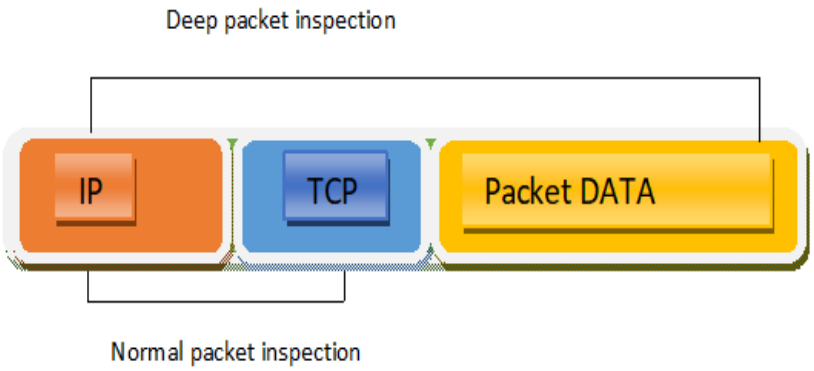

Figure-1: Deep Packet Inspection

These extracted results from the raw network dataset used for further processing.

2.3.2. Preprocessing

Once the features are extracted via DPI, we follow several steps to prepare the data for machine learning or deep learning models:

Step 1: Handling Missing or Invalid Values by remove or impute missing values with the mean, median or mode of the column.

Step 2: If categorical features exist, perform Feature Encoding by converting categorical data (like protocol type: TCP, UDP, ICMP) into numerical format.

- **Label Encoding**: TCP → 1, UDP → 2, ICMP → 3
- **One-Hot Encoding**: Create binary columns for each protocol

Step 3: Standardization is doing to shift the data so that the average become zero and scale the data according to the variance. It is especially useful when features have different units.

$$X_{standardized}=\frac{X-\mu}{\sigma}$$

Where:

- X - is the original value
- M - is the mean of the feature
- σ - is the standard deviation of the feature

Result: After normalization the dataset shows a zero mean and unit variance.

Step 4: Normalization is used to scale the data into a specific range [0, 1]. This type of pre-processing is useful for standardization and it is important for algorithms that depend on the magnitude of the data.

Min - Max Normalization

$$X_{normalized}=\frac{X-Xmax}{Xmax-Xmin}$$

Where:

- X - is the original value
- X_{min} and X_{max}: The minimum and maximum values of the feature

Result: All feature values are scaled between 0 and 1

2.3.3. Final Output

After preprocessing, the data becomes clean, consistent, and scaled and ready for input into the RFE to eliminate the less important data and thus we can improve the time in the phase of learning algorithms.

2.4 Recursive Feature Elimination (RFE)

To reduce dimensionality and enhance training speed, RFE iteratively removes low-impact features. This improves the focus of the learning algorithm on influential attributes [9].

After preprocessing, the dataset becomes clean, consistent, and appropriately scaled, making it suitable for input into feature selection techniques such as RFE. This method is particularly useful for reducing dimensionality, enhancing model interpretability, minimizing over fitting, and reducing the time complexity of training machine learning and deep learning models. Recursive Feature Elimination (RFE) is applied as a model-driven feature selection technique. It iteratively trains a learning model and discards the least significant features based on their

importance scores. This process continues until the target number of features is reached, retaining only the most informative ones for training.

The major idea is to eliminate irrelevant or redundant features that do not contribute significantly to the models predictive power.

Working Mechanism of RFE

Let:

$X \in R^{nxd}$ be the input data after preprocessing, where:

- n: number of samples
- d: number of features

$y \in R^{n}$ be the target vector

M be a base estimator

1. **Model Training:** A learning model M is trained on the current feature set $X^{(t)} \in R^{n \times d}$, where d_t is the number of features at iteration t.

$\hat{y}(t)=M^{(t)}(X^{(t)})$

2. **Feature Importance Calculation:** In the case of linear models, feature importance is evaluated using the absolute values of the model's learned coefficients. Features with smaller magnitudes are deemed less significant and are gradually removed through the RFE process.

$I_j^{(t)}=|w_j^{(t)}| \text{ for } j=1,2,\ldots,d_t$

where $w_j^{(t)}$ is the weight or coefficient assigned to the j-th feature.

For tree-based models, importance scores may be based on impurity reduction.

3. **Feature Elimination:** Identify the feature with the lowest importance score:

$j_{min}^{(t)}=arg\ min_j\ I_j^{(t)}$

Remove this feature from the dataset:

$X^{(t+1)}=X^{(t)} \setminus \{ j_{min}^{(t)} \}$

4. **Repeat:** Steps 1 to 3 are repeated until the number of remaining features d_t equals the predefined number of desired features k, i.e., $d_t=k$

Final Output

After RFE completes, the resulting feature matrix $X^{(T)} \in R^{n \times k}$ contains only the most informative features. These are then passed into downstream machine learning or deep learning models for training and evaluation. This structured elimination process leads to a more compact and efficient representation of the input data, which not only enhances the model's learning performance but also reduces computational complexity.

2.5 Variational Autoencoders(VAE)

After the dimensionality reduction via RFE, the reduced and refined feature set is well-suited for deep generative learning models like Variational Autoencoders (VAE), which can capture complex patterns of network traffic for distinguishing the normal and malicious behaviors [10].

A VAE is a type of autoencoder designed to reconstruct input data and also to learn a latent distribution of the input [11]. In DDoS detection, this allows the system to model the normal behavior of network traffic, making it easier to flag anomalies that deviate significantly from this learned distribution.

- $X \in R^{n \times k}$ be the input after RFE (reduced to k features)
- $Z \in R^{n \times l}$ be the latent variable space (with l<k)

The encoder transforms the input data X into a probability distribution within the latent space.

$q_\phi(z|x)=(N(z;\mu(x),\sigma^2(x)))$

Where:

- Φ - are the encoder parameters
- $\mu(x)$ and $\sigma(x)$ - are the mean and variance vectors predicted by the encoder network

Using the latent representation z, the decoder generates a reconstruction of the original input data.

$p_\theta(x|z)$

The VAE objective is to minimize the Evidence Lower Bound (ELBO):

$L(\theta,\phi;x)=E_{q\phi(z|x)}[\log p\theta(x|z)]-D_{KL}(q_\phi(z|x)||p(z))$

Where:

- The first term is the reconstruction loss

- The second term is the Kullback–Leibler divergence that regularizes the encoder by making the approximate posterior $q_\phi(z/x)$ close to a prior $p(z) \sim N(0,I)$

In DDoS detection, VAE can be trained on normal traffic. During inference, inputs with high reconstruction errors are flagged as potential anomalies or attacks.

2.6 Attention-Enriched Transfer Learning (AETL)

Attention-Enriched Transfer Learning (AETL) enhances traditional transfer learning by incorporating attention mechanisms that focus the model's capacity on the most relevant features in the input. This is especially critical for multi-source, high-dimensional data such as network flows involved in DDoS detection [12].

Transfer Learning Backbone: A pre-trained model, Transformer encoder, is used, trained on a large dataset (like CICIDS2019, CICIDS2020, and TON_IoT) and fine-tuned on the task-specific DDoS data. **Attention Module:** Attention layers are integrated to learn feature weights dynamically, focusing on the most important time-steps or input features.

Let the attention score for feature i at time ttt be:

$\alpha_{t,i} = \frac{exp\,(et,i)}{\sum j=1 kexp\,(et,i)}$

where $e_{t,i}$=score(h_t,x_i)

Where h_t is the hidden state or context vector and x_i is the input feature.

Weighted Representation: The final representation is a weighted sum:

$$x'_t = \sum_{i=1}^{k} \alpha_{t,i} \cdot x_{t,i}$$

Fine-Tuning: The full model (backbone + attention) is fine-tuned on the labeled DDoS dataset.

2.7 .LSTM for Temporal Correlation

LSTM networks are employed to detect long-term dependencies in packet sequences. This is especially useful in identifying stealthy low-rate DDoS attacks [13].

To capture temporal dependencies in sequential network traffic data, LSTM networks are highly effective. LSTM is a type of Recurrent Neural Network (RNN) capable of learning long-term dependencies, making it ideal for detecting gradual or stealthy DDoS attacks over time[14].

Each LSTM cell maintains a cell state - C_t and hidden state - h_t , updated using gates:

1. **Forget Gate:**

$f_t=\sigma(W_f \cdot [h_{t-1},x_t]+b_f)$

2. **Input Gate:**

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad , \quad \tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

3. **Update Cell State:**

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

4. **Output Gate:**

$o_t=\sigma(W_o \cdot [h_{t-1},x_t]+b_o)$, $h_t=o_t \cdot tanh(C_t)$

Where:

- x_t - input at time t
- σ - sigmoid activation
- tanh - hyperbolic tangent activation

Models sequential behavior of traffic (e.g., bursts, repeated pings, time-dependent anomalies). Learns both short-term fluctuations and long-term trends. Can be trained on raw or RFE-selected input sequences

After applying RFE, VAE learns latent representations of normal traffic for unsupervised anomaly detection. AETL leverages transfer learning with attention to detect patterns even with limited data. LSTM models the sequential nature of DDoS attacks, capturing both fast and slow attack behaviors.

These models can be used independently or in ensemble/hybrid form to enhance DDoS detection effectiveness in terms of accuracy, precision, recall, and response time.

2.8 The proposed Model Architecture.

The proposed combined hybrid model leverages the strengths of feature engineering, generative modelling, transfer learning, and sequence modelling to robustly detect DDoS attacks across various types and traffic patterns, improving both accuracy and generalization. The proposed new hybrid algorithm is follows:

$$\hat{y}=\sigma(W_o \cdot \tanh(C_T) \cdot softmax(score(h_T,x)) \cdot f_{VAE}(f_{RFE}(f_{norm}(f_{std}(f_{enc}(f_{miss}(X)))))+b_o)$$

where:

Input (X): Features are extracted from raw network traffic using Deep Packet Inspection (DPI).

- Missing values are handled (f_{miss})
- Categorical data is encoded (f_{enc})
- Features are standardized (f_{std}) and normalized (f_{norm})for consistent scale.
- f_{RFE} selects the most important features to reduce dimensionality and improve learning efficiency.
- f_{VAE} transforms the reduced feature set into a latent space that captures complex, non-linear patterns, enhancing anomaly detection.
- Attention *(softmax(score))* highlights the classification tasks most relevant input features or time steps .
- The LSTM captures temporal patterns in traffic sequences, helping detect short-term spikes and long-term stealthy attacks.
- C_T and h_T represent the final cell and hidden states.

The output is passed through a sigmoid activation σ, yielding $\hat{y}$, the probability that the input represents a DDoS attack.

3. EXPERIMENTAL SETUP

3.1 Datasets

1 We use the CIC-DDoS2020 Dataset, which reflects recent DDoS attack scenarios[15].
2. We use the CIC-DDoS2019 and TON_IoT datasets, which reflects recent DDoS attack scenarios [16], [17].
3. We use the CIC-DDoS2017 Dataset, which reflects recent DDoS attack scenarios

3.2 Evaluation Metrics

The model is evaluated using Accuracy, Precision, Recall, F1-Score, ROC-AUC, and Detection Time.

4. RESULTS AND ANALYSIS

4.1 Performance Comparison

Table 2 compares the proposed model with dataset CIC-DDoS2019 with other machine learning algorithms such as CNN-LSTM, Random forest, SVM and LSTM

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Detection Time (ms)
Proposed Hybrid Model	**98.45**	**98.33**	**98.56**	**98.44**	**5.2**
CNN-LSTM	97.23	97.81	97.12	97.96	6.0
Random Forest	95.78	95.32	95.93	95.62	7.1
SVM	91.74	91.23	91.80	91.52	8.5
Traditional LSTM	94.81	94.14	94.94	94.54	7.0

Table 2. Performance analysis of algorithms with dataset CIC-DDoS2019

Table 3 compares the proposed model with dataset CIC-DDoS2020 with other machine learning algorithms such as CNN-LSTM, Random forest, and LSTM

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Detection Time (ms)
Proposed Hybrid Model	**98.00**	**98.50**	**97.71**	**98.60**	**5.0**
CNN-LSTM	97.14	97.89	97.35	98.12	6.1
Random Forest	95.65	95.27	95.91	95.59	7.8
Traditional LSTM	94.84	94.23	95.01	94.62	6.7

Table 3. Performance analysis of algorithms with dataset CIC-DDoS2020

Table 4 compares the proposed model with dataset TON_IoT with other machine learning algorithms such as CNN-LSTM, Random forest, SVM and LSTM

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Detection Time (ms)
Proposed Hybrid Model	**97.92**	**97.76**	**98.02**	**98.89**	**5.3**
CNN-LSTM	96.67	96.44	96.89	96.66	6.3
Random Forest	94.21	93.92	94.13	94.02	7.2
SVM	90.68	90.21	90.72	90.46	8.4
Traditional LSTM	93.45	93.12	93.55	93.33	7.3

Table 4. Performance analysis of algorithms with dataset TON_IoT

Table 5 compares the proposed model with dataset CIC-DDoS2017 with our previous hybrid model and the proposed model.

Algorithm	Accuracy	Precision	Recall	F1 Score	Detection Time (ms)
Hybrid model with Q-learning	96.2%	95.8%	96.0%	95.9	1.9
The proposed hybrid model with VAE	98.01	97.2	98.2	97.9	1.5

Table 5. Performance analysis of algorithms with dataset CIC-DDoS2017

4.2 Hypothetical Comparison Graph:

Here is the hypothetical bar chart illustrating the performance of these algorithms showing, the superiority of the proposed hybrid model, which has a high detection speed and higher accuracy.

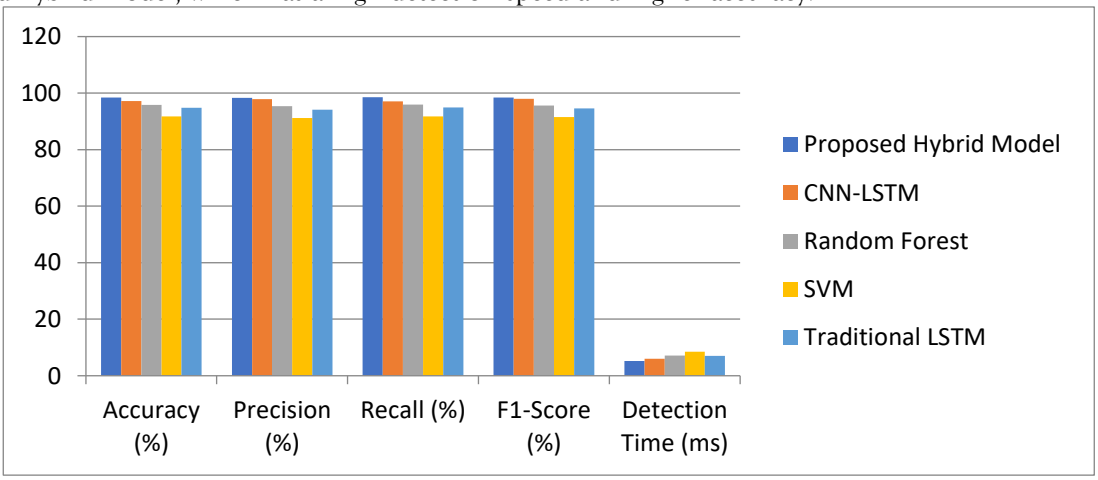

Figure-2: Performance Analysis with CIC-DDoS2019 dataset

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 20s, 2025
https://www.theaspd.com/ijes.php

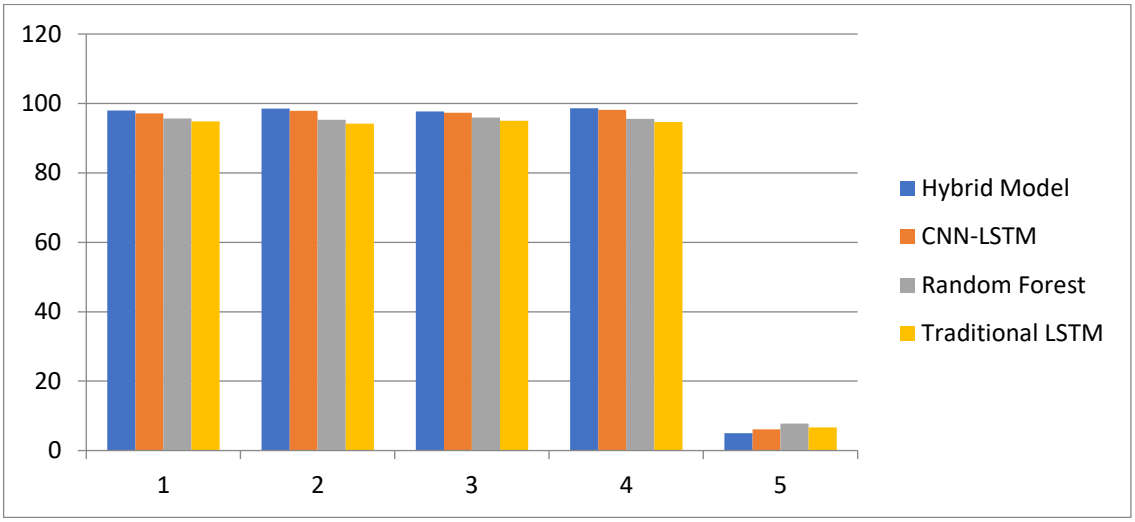

Figure-3: Performance Analysis with CIC-DDoS2020 dataset

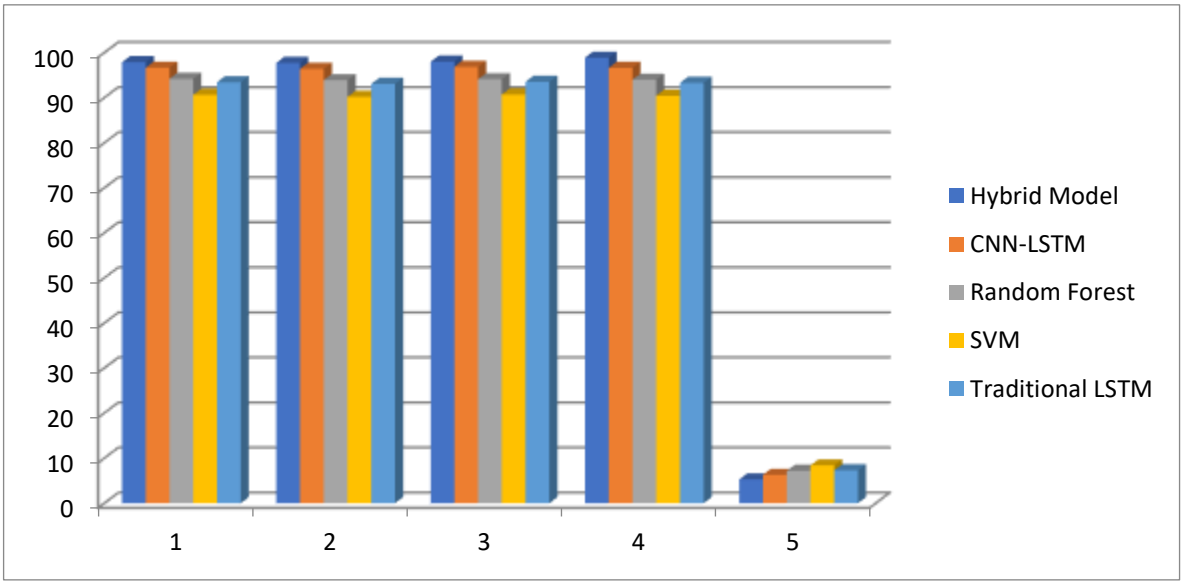

Figure-4: Performance Analysis with TON IoT dataset

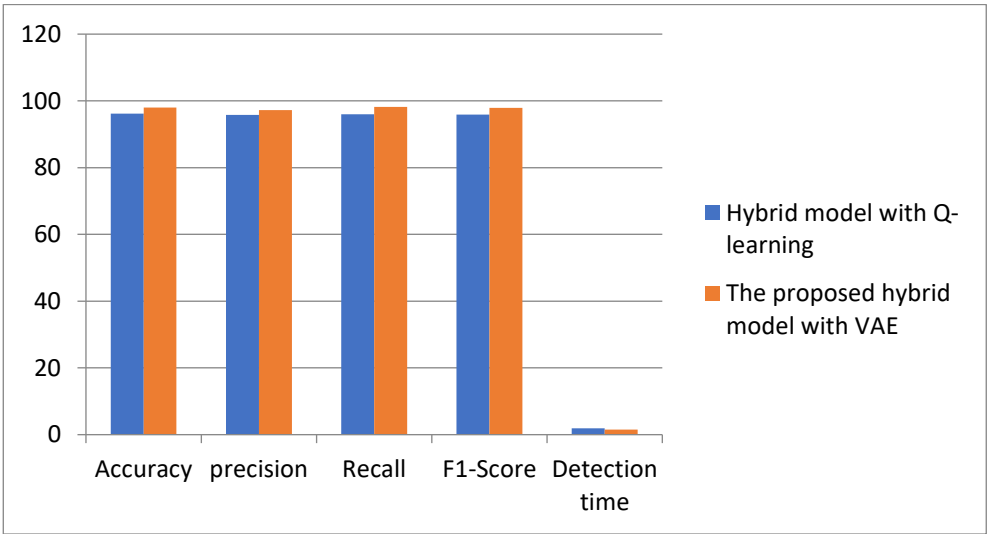

Figure-4: Performance Analysis with CIC-DDoS2017 dataset

4.2 Analysis:

The Next-Gen hybrid DDoS detection model shows a better performance across with multiple datasets as well as with the some of the major algorithms. By combining Deep Packet Inspection, Recursive Feature elimination, variational Auto-encoders, attention enriched transfer learning, and Long Short-Term Memory, the system can capture the major features present across the network ,learns it and thus it can detect both known and unknown attack patterns. The VAE helps the system to capture high-dimensional data. The AETL helps the model to focus only on relevant features and thus improves the generalization among the datasets.

The proposed model magnifies the system's ability to distribute various forms of attacks with high precision, recall and accuracy. The above comparison shows a clear furtherance over the earlier models. The experimental results with CIC-DDoS2019, CIC-DDoS2020, TON_IoT and CIC-DDoS2017 datasets recommends the models ability , strong performance and the adaptability.

The models cohesive decision Layer (VAE, AETL, LSTM) contributes to balance the strengths and reimburse the weakness of each component and thus results a trusted system to detect the DDoS. There for the model is an effective, reliable and adaptable solution for DDoS detection in heterogeneous environments.

5. CONCLUSION AND FUTURE WORK:

The proposed Next-Gen DDoS detection system shows a clear improvements among both traditional and existing detection models. Here we combined Deep Packet Inspection, Recursive Feature elimination, Variational Auto-encoders, attention enriched transfer learning, and Long Short-Term Memory and thus the model can detect complex patterns among different network traffic. The VAE helps to detect both trained and non trained patterns and attacks by learning the normal pattern of the network traffic. AETL and LSTM helps in filtering the features by analyzing the traffic behavior. The experimental results with CIC-DDoS2019, CIC-DDoS2020, TON_IoT and CIC-DDoS2017 datasets recommends the models ability, strong performance and the adaptability.

Overall, the outcome of this work results in a flexible DDos Detection Framework , which can work with enterprise as well as with IOT-Based environments. Its strong performance across multiple datasets shows that it can adapt to different types of network conditions.

In the future, our focus will be on optimizing the system for real-time use in IoT networks. This includes making the model faster and more energy-efficient for use on edge devices, and exploring new techniques like adversarial learning to improve its defense against advanced attacks. These steps will help make the system even more suitable for real-world IoT applications, where quick, accurate, and lightweight solutions are essential.

6. REFERENCES:

1. K. Al-Begain, M. Khan, B. Alothman, C. Joumaa, and E. Alrashed, "A DDoS Detection and Prevention System for IoT Devices and Its Application to Smart Home Environment," *Applied Sciences*, vol. 12, no. 22, p. 11853, Nov. 2022, doi: https://doi.org/10.3390/app122211853.
2. S. Sadhwani, Baranidharan Manibalan, Raja Muthalagu, and P. M. Pawar, "A Lightweight Model for DDoS Attack Detection Using Machine Learning Techniques," *Applied sciences*, vol. 13, no. 17, pp. 9937–9937, Sep. 2023, doi: https://doi.org/10.3390/app13179937.
3. K.-H. Le, M.-H. Nguyen, T.-D. Tran, and N.-D. Tran, "IMIDS: An Intelligent Intrusion Detection System against Cyber Threats in IoT," *Electronics*, vol. 11, no. 4, p. 524, Feb. 2022, doi: https://doi.org/10.3390/electronics11040524.
4. M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *Journal of Network and Computer Applications*, vol. 60, pp. 19–31, 2016
5. Z. Fan, Y. Wang, L. Meng, G. Zhang, Y. Qin, and B. Tang, "Unsupervised Anomaly Detection Method for Bearing Based on VAE-GAN and Time-Series Data Correlation Enhancement (June 2023)," *IEEE Sensors Journal*, vol. 23, no. 23, pp. 29345–29356, Oct. 2023, doi: https://doi.org/10.1109/jsen.2023.3326335.
6. M. Rich, R. Mills, T. Dube, and S. Rogers, "Evaluating Machine Learning Classifiers for Defensive Cyber Operations," *Military Cyber Affairs*, vol. 2, no. 1, Dec. 2016, doi: https://doi.org/10.5038/2378-0789.2.1.1005.
7. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: https://doi.org/10.1145/1541880.1541882.
8. B. K. A and M. Vijayakumar, "Enhancing Intrusion Detection System (IDS) Through Deep Packet Inspection (DPI) with Machine Learning approaches," Apr. 2024, doi: https://doi.org/10.1109/adics58448.2024.10533473.
9. X. Zeng, Y.-W. Chen, and C. Tao, "Feature Selection Using Recursive Feature Elimination for Handwritten Digit Recognition," Sep. 2009, doi: https://doi.org/10.1109/iih-msp.2009.145.
10. D. Kingma and M. Welling, "Auto-Encoding Variational Bayes," 2014. Available: https://arxiv.org/pdf/1312.6114

11. H. Wang, J. Yan, and N. Jia, "A New Encrypted Traffic Identification Model Based on VAE-LSTM-DRN," *Computers, materials & continua/Computers, materials & continua (Print)*, vol. 78, no. 1, pp. 569–588, Jan. 2024, doi: https://doi.org/10.32604/cmc.2023.046055.
12. A. Vaswani *et al.*, "Attention Is All You Need," *arXiv*, Jun. 12, 2017. https://arxiv.org/abs/1706.03762
13. S. A. Althubiti, E. M. Jones, and K. Roy, "LSTM for Anomaly-Based Network Intrusion Detection," *IEEE Xplore*, Nov. 01, 2018. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8615300 (accessed Apr. 04, 2022)
14. M. A. Ferrag *et al.*, "Generative AI in Cybersecurity: A Comprehensive Review of LLM Applications and Vulnerabilities," *Internet of Things and Cyber-Physical Systems*, Feb. 2025, doi: https://doi.org/10.1016/j.iotcps.2025.01.001.
15. "DDoS 2019 | Datasets | Research | Canadian Institute for Cybersecurity | UNB," *www.unb.ca*. https://www.unb.ca/cic/datasets/ddos-2019.html
16. "DDoS 2019 | Datasets | Research | Canadian Institute for Cybersecurity | UNB," *www.unb.ca*. https://www.unb.ca/cic/datasets/ddos-2019.html
17. A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, and A. Anwar, "TON_IoT Telemetry Dataset: A New Generation Dataset of IoT and IIoT for Data-driven Intrusion Detection Systems," *IEEE Access*, vol. 8, pp. 1–1, 2020, doi: https://doi.org/10.1109/access.2020.3022862.
18. A. A. Alahmadi *et al.*, "DDoS Attack Detection in IoT-Based Networks Using Machine Learning Models: A Survey and Research Directions," *Electronics*, vol. 12, no. 14, p. 3103, Jan. 2023, doi: https://doi.org/10.3390/electronics12143103.
19. Boopathi Chettiagounder Sengodan *et al.*, "Variational Autoencoders for Network Lifetime Enhancement in Wireless Sensors," *Sensors*, vol. 24, no. 17, pp. 5630–5630, Aug. 2024, doi: https://doi.org/10.3390/s24175630.
20. Y. Yuan, Z. Chen, Z. Wang, Y. Sun, and Y. Chen, "Attention mechanism-based transfer learning model for day-ahead energy demand forecasting of shopping mall buildings," *Energy*, vol. 270, p. 126878, May 2023, doi: https://doi.org/10.1016/j.energy.2023.126878.
21. GIBI K S and DR.S.NITHYA, "Hybrid Detection Model Combining an Advanced Q-Learning Network (AQN) and Attention-Enriched Transfer Learning Approach for DDoS Detection," *Nanotechnology Perceptions*, pp. 246–254, Dec. 2024, doi: https://doi.org/10.62441/nano-ntp.vi.3708.