



A Deep approach For Breach Detection Using Temporal Fusion Transformers

¹Vajrala.Siddhardhareddy, ²Rayana.Madhu, ³Rachamanti.SaiViswanath,

⁴Rachakonda.Santhosh, ⁵Mr.Yeriniti.Venkata Narayana

^{1,2,3,4}Dept. of Information Technology, VVIT, Andhra Pradesh, India

⁵Assistant Professor. Dept. of Information Technology, VVIT, Andhra Pradesh, India

Abstract: Breach detection helps in identifying unauthorized access or suspicious activities in a network system. It detects threats at an early stage before they cause serious damage, protecting sensitive information like personal data, financial records, and confidential files. Breach detection systems provide real-time alerts, allowing security teams to take quick action and prevent further harm. However, most existing breach detection systems face challenges like false alarms, slow detection speed, and inability to detect new attacks. Some systems also struggle to detect intrusions in encrypted data and consume high system resources. These limitations affect the accuracy and performance of the detection system. To overcome these issues, the breach detection system is implemented using Temporal Fusion Transformers (TFT), which analyses time-based patterns in network traffic to detect intrusions accurately. The current study incorporates the Simargyl2022 dataset to enhance the quality of our results and analyses, which contains both normal network traffic and malicious attack data, making it suitable for evaluating detection performance. The system achieved 95.40 accuracy, with a recall of 95.40, precision of 91.01, and an F1-score of 93.15, showing its high efficiency in detecting breaches. The outcomes of this study have significant implications for network security, providing valuable insights for practitioners and researchers working towards building robust and intelligent breach detection systems.

Key Words: Breach Detection, Temporal Fusion Transformer (TFT), Cybersecurity, Anomaly Detection, Time-Series Forecasting, Temporal Dependencies, Multi-Headed Attention, Gating Layers, Scalable Systems, Advanced Deep Learning.

1. INTRODUCTION

In the modern technological era, the rapid advancement of Information and Communication Technology (ICT) has led to the widespread adoption of digital systems in various organizations. This integration has resulted in a growing reliance on network-based systems, making organizations increasingly vulnerable to cyberattack. As the complexity and frequency of cyber threats continue to increase, However, traditional security measures are inadequate. Therefore, the need for an advanced breach detection system that can not only identify known attacks but also adapt to new and evolving threats have become a critical necessity for cybersecurity.

Breach detection plays a vital role in the security of network infrastructures by identifying unauthorized access, suspicious ac- levels, and malicious intrusions. Existing cybersecurity methods, such as Intrusion Detection Systems (IDS) and anomaly detection techniques are widely used to protect networks. However, anomaly detection techniques face challenges in defining complex rules and often fail to recognize harmful activities that resemble normal patterns. In contrast, IDSs have proven to be more effective in analysing network traffic and predicting potential security breaches based on the trained algorithms. However, traditional IDSs still struggle with high false alarm rates and lack adaptability to novel attack patterns.

Recent advancements in deep learning have significantly enhances the capabilities of cybersecurity systems. Deep learning, a specialized subset of machine learning, mimics human brain function by using artificial neural networks. These networks can automatically learn complex patterns from data, making them par- are highly effective for intrusion detection. Deep learning models such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and Recurrent Neural Networks (RNN) have been applied in cybersecurity with promising results. However, these models still require improvements in handling time-series data and real-time security breach.

To address these challenges, this research introduces a Deep Approach for Breach Detection using Temporal Fusion Transformers (TFT). The TFT model is particularly well-suited for analysing time- dependent patterns in network traffic,



allowing it to detect both known and unknown cyberthreats. By leveraging sequential data processing and attention mechanisms, TFT enhances the ability to identify complex cyberattacks that evolve over time. The system is trained and tested using the Simargyl2022 dataset, which includes a diverse range of normal network traffic and malicious activity patterns. This ensures the model's robustness and adaptability to Real-world cybersecurity challenges.

TFT is a powerful deep learning model designed to handle complex time-series forecasting tasks by efficiently capturing long-range dependencies. Due to its interpretability and predictive accuracy, it has gained traction across multiple industries. This figure highlights the widespread adoption of TFT in domains such as finance, healthcare, cybersecurity, and environmental monitoring, demonstrating its versatility in solving real-world problems.

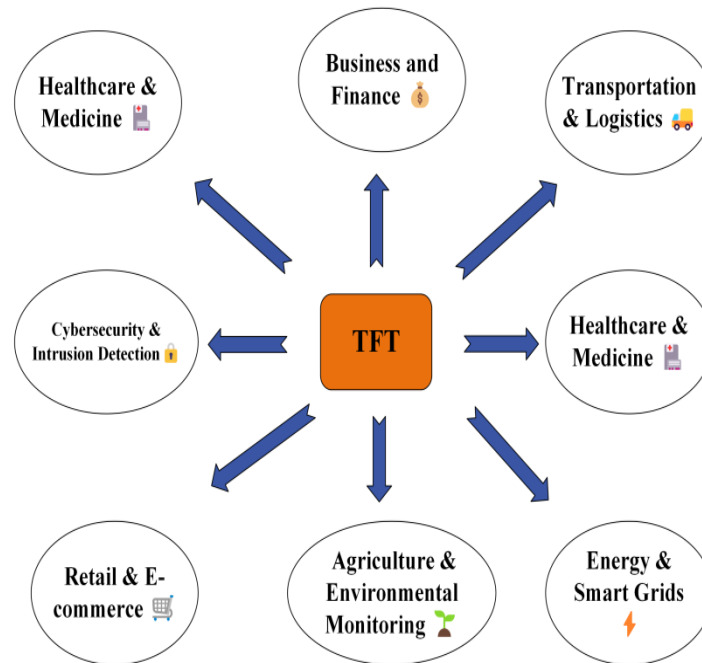


Fig-1: Applications Of TFT

2. LITERATURE REVIEW

Vigneswaran et al. (2018) evaluated the performance of both shallow and deep neural networks for network intrusion detection. Their study compared various architectures, including fully connected deep neural networks (DNNs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs). The results highlighted that deep learning models significantly outperform traditional machine learning algorithms in detecting network intrusions. However, they also noted challenges such as high computational costs and the need for extensive hyperparameter tuning.

Zhang et al. (2020) investigated deep learning models for anomaly detection in network traffic. The study compared CNNs, LSTMs, and hybrid deep learning models. Their findings indicated that while CNNs are effective at feature extraction, LSTMs and attention-based mechanisms (such as those in Transformers) improve sequence modeling for real-time intrusion detection. The study highlighted the need for models that can adapt to evolving attack strategies, reinforcing the importance of TFT in breach detection systems.

Karim et al. (2021) explored the impact of deep temporal learning models in cybersecurity. The study focused on the ability of sequence-based models such as LSTMs and Transformers to capture complex attack patterns over time. The authors concluded that models leveraging temporal dependencies, such as the Transformer architecture, outperform traditional methods in identifying sophisticated cyber threats. This aligns with the motivation behind using **Temporal Fusion Transformers (TFT)** in breach detection, as TFT effectively integrates past information while capturing dynamic patterns in network traffic.

Lim et al. (2021) introduced the **Temporal Fusion Transformer (TFT)** as an advanced deep learning model designed for interpretable time-series forecasting. Their study demonstrated how TFT could integrate static and dynamic features while maintaining temporal dependencies. The model's ability to handle mixed-data types and long-range dependencies



makes it an ideal choice for anomaly detection in cybersecurity. The findings suggest that TFT's attention mechanism enhances interpretability, allowing security analysts to understand why certain breaches are flagged.

Vanin et al. (2022) provided an extensive survey of artificial intelligence (AI) and machine learning (ML) approaches for intrusion detection. The study emphasized the effectiveness of ensemble methods and deep learning techniques in identifying cyber threats. The authors noted that while traditional machine learning methods such as Support Vector Machines (SVM) and Decision Trees perform well on smaller datasets, deep learning approaches like Long Short-Term Memory (LSTM) networks and Transformer-based models excel in handling large-scale intrusion datasets. Their findings suggest that deep models, particularly those integrating temporal patterns, can enhance intrusion detection accuracy.

Huang et al. (2022) explored the use of Transformer-based architectures in cybersecurity applications. Their research focused on using self-attention mechanisms to analyze network traffic logs and detect anomalies. The study found that Transformers outperform conventional RNN-based models by reducing vanishing gradient issues and improving long-term dependency learning. The results support the integration of Temporal Fusion Transformers for network breach detection, as they can efficiently model sequential dependencies while enhancing detection accuracy.

3. METHODOLOGY

3.1 Data Collection

The proposed methodology commences with the collection of a comprehensive dataset encompassing a diverse array of network traffic instances, which include both routine activities and various cyber threats such as network spoofing, Denial of Service (DoS), and Distributed Denial of Service (DDoS) attacks. This dataset serves as the foundation for the training and evaluation of breach detection systems.

3.2 Data Preprocessing

Following the acquisition of the dataset, preprocessing procedures were implemented to ensure the data's quality and consistency. The network traffic data were imported into memory utilizing the Pandas library, and feature extraction was performed to identify key attributes such as packet size, duration, protocol type, and the IP addresses of both the source and destination. Subsequently, the input data were meticulously aligned to ensure consistent feature representation.

3.3 Data Augmentation

To enhance the model's robustness, data augmentation techniques were employed. This process involves the elimination of attributes that do not significantly contribute to intrusion detection. By removing redundant or irrelevant features, the model's performance can be improved while also reducing computational demands.

3.4 Data Encoding

Given that machine learning and deep learning models predominantly operate with numerical data, categorical features in the dataset were transformed using one-hot encoding. This process converts categorical values, such as protocol types and IP addresses, into a numerical format suitable for model training.

3.5 Data Splitting

Subsequent to preprocessing, the dataset is partitioned into training and testing subsets using the `train_test_split` function from Scikit-learn. Typically, 80% of the data is allocated for training, while the remaining 20% is reserved for testing. This approach ensures the model is trained effectively and evaluated on data it has not previously encountered, which is crucial for assessing the model's generalization capability.

Model Architecture Design

Temporal Fusion Transformer (TFT):

The Temporal Fusion Transformer (TFT) is a deep learning model specifically engineered for time-series forecasting and sequential data analysis. It incorporates multiple mechanisms, including attention mechanisms, gating layers, and feature selection, which enhance its efficacy in network traffic analysis for intrusion detection. In contrast to traditional deep learning models such as Long Short-Term Memory networks (LSTMs) or Convolutional Neural Networks (CNNs), the



TFT is adept at capturing both long-term dependencies and short-term patterns in network data, rendering it suitable for identifying cyber threats that evolve over time.

In the described system, the TFT model is trained using network traffic datasets to classify normal and malicious activities. The dataset undergoes preprocessing, label encoding, and division into training and testing sets prior to being input into the TFT model. Once trained, the model extracts pertinent features from network traffic data, enabling it to identify anomalies and detect potential intrusions with high accuracy. Its capacity to manage complex dependencies in sequential data positions it as a powerful tool in contemporary intrusion detection systems.

Furthermore, the TFT's interpretability provides valuable insights into the contribution of different network traffic features toward classification, making it a more transparent and explainable AI-driven intrusion detection approach. Its ability to adapt to varying network conditions and evolving cyber threats makes it an effective and scalable solution for contemporary cybersecurity frameworks. By leveraging historical and real-time network traffic data, the TFT-based system can proactively detect sophisticated attacks, including zero-day exploits, denial-of-service (DoS) attacks, and advanced persistent threats (APTs), thereby strengthening overall network security and resilience.

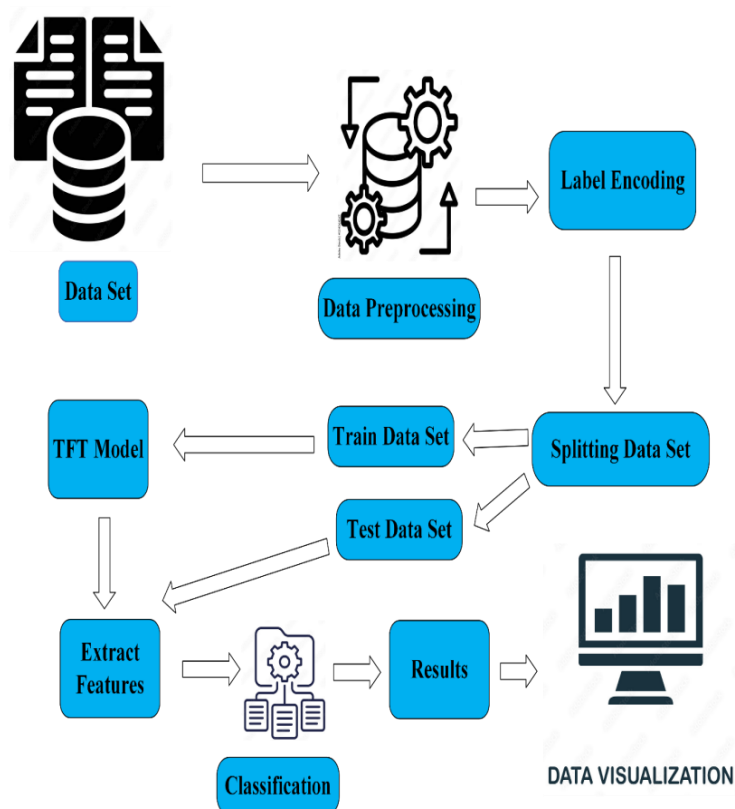


Fig-2: System Architecture

Model Training and Evaluation

Once the Temporal Fusion Transformer (TFT) model is structured, it undergoes training using the processed dataset from Simargl2022. The primary objective of this training phase is to enable the model to accurately classify network traffic as either normal or an intrusion attempt. The training and evaluation procedure follows these key steps:

Model Compilation

The model is initialized with appropriate hyperparameters to optimize performance for intrusion detection. Unlike conventional models, TFT does not rely on explicit optimizers in our approach, making it uniquely suited for learning temporal dependencies within network traffic. Instead, it utilizes its gating mechanisms and variable selection networks to improve efficiency.



Training Process

The TFT model is trained using the pre-processed network traffic data, where each instance contains both numerical and categorical features. The training runs for a set number of epochs, ensuring the model learns intricate patterns of cyber threats. The data is fed into the model in mini-batches to enhance computational efficiency and prevent overfitting.

Evaluation

Once the training process is complete, the model's effectiveness is assessed using the testing dataset. Key performance metrics, including accuracy, precision, recall, and F1-score, are calculated to evaluate its capability in identifying and mitigating network intrusions.

Real-Time Deployment and Evaluation

After completing the training and assessment phases, the developed model is implemented in a live network environment to actively monitor traffic and identify potential intrusions.

System Integration

The trained model is incorporated into the existing network framework, enabling it to analyse real-time traffic patterns and detect suspicious activities efficiently.

Performance Testing

Rigorous testing and validation procedures are conducted to measure the system's effectiveness in practical scenarios. The model's ability to recognize and mitigate different types of cyber threats is analysed, and necessary refinements are applied to enhance overall performance.

Performance Evaluation

After training, the model is evaluated using the test dataset to assess its ability to detect intrusions. Key performance metrics such as accuracy, precision, recall, and F1-score are computed to determine the model's effectiveness. Additionally, the interpretability features of TFT allow for insight into which network attributes contribute most to anomaly detection, helping refine the intrusion detection process.

Key Features of TFT in Proposed System:

1. **Multi-Horizon Forecasting:** TFT predicts future network states by analysing past traffic data, helping identify evolving cyber threats.
2. **Attention Mechanism:** The model assigns different importance levels to various time steps, ensuring that critical network behaviours are prioritized for detection.
3. **Gating Layers:** These layers control information flow, preventing irrelevant data from affecting intrusion detection accuracy.
4. **Feature Selection:** TFT identifies the most relevant features in network traffic, reducing noise and improving model performance.
5. **Handling Imbalanced Data:** Cybersecurity datasets often contain a small number of attack samples. TFT effectively manages this imbalance, ensuring rare intrusions are detected without being overshadowed by normal traffic.

Advantages of the Proposed System:

1.Improved Detection Accuracy

The integration of the Temporal Fusion Transformer (TFT) enhances the detection of cyber threats by effectively capturing long-term dependencies and temporal patterns in network traffic data. This results in higher accuracy in identifying both known and emerging attacks.



2. Reduction in False Positives

Traditional intrusion detection systems often suffer from high false positive rates. The proposed system leverages TFT's attention mechanisms and feature selection capabilities, reducing misclassifications and improving reliability.

3. Adaptive and Scalable

Unlike conventional machine learning models, the TFT model is highly adaptive to dynamic network environments. It can analyse real-time data, making it suitable for large-scale cybersecurity applications with evolving attack patterns.

4. Effective Feature Extraction

The system automatically extracts relevant features from network traffic, reducing the need for extensive manual feature engineering. This ensures better learning and improved performance for detecting anomalies.

5. Enhanced Interpretability

TFT provides explainable predictions by highlighting the key features influencing its decisions. This interpretability is critical in cybersecurity, as it allows security analysts to understand and mitigate threats effectively.

6. Robust Against Imbalanced Data

Many cybersecurity datasets have a higher proportion of normal traffic compared to attack traffic. TFT's ability to handle imbalanced data ensures that minority attack classes are detected effectively without being overshadowed by normal traffic.

7. Real-Time Threat Detection

The proposed system processes network traffic efficiently, enabling near real-time detection of intrusions. This allows for quicker responses to security threats, reducing potential damage.

By leveraging the Temporal Fusion Transformer, the proposed system significantly enhances breach detection capabilities, making it a powerful solution for modern cybersecurity challenges.

RESULTS

Performance Analysis

In Table 1, the evaluation metrics for the breach detection system using the Temporal Fusion Transformer (TFT) on the Simargyl2022 dataset demonstrate strong performance. Key metrics such as Accuracy, Precision, Recall, and F1-score are essential in assessing the model's effectiveness in detecting security breaches. The results highlight the TFT model's superior ability to accurately classify network traffic as either normal or intrusive, emphasizing its robustness in identifying cyber threats.

Performance Metrics

The following metrics are used to assess the model's effectiveness:

5.1 Accuracy :

Accuracy is one of the most commonly used evaluation metrics for classification models, including Threat Detection Systems (TDS).

It measures how often the model makes the correct predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots\dots(1)$$

Where:

- TP (True Positives): Correctly predicted threats.
- TN (True Negatives): Correctly predicted benign traffic.



- FP (False Positives): Incorrectly classified benign traffic as a threat (false alarm).
- FN (False Negatives): Incorrectly classified threats as benign (missed attacks).

5.2 Precision

Precision (also known as Positive Predictive Value) measures how many of the predicted positive cases (threats) were actually correct. It is crucial in Threat Detection Systems (TDS) to minimize false positives (FP) and avoid unnecessary security alerts.

Precision Formula

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots(2)$$

Where:

TP (True Positives): Correctly predicted threats.

FP (False Positives): Incorrectly classified benign traffic as threats (false alarms).

5.3 F1 Score:

The **F1-score** is a performance metric used to evaluate the balance between **precision** and **recall** in classification models. It is particularly useful when dealing with imbalanced datasets, where focusing solely on accuracy can be misleading.

The F1-score is the harmonic mean of precision and recall,

calculated as:

$$F1 = \frac{Precision \times Recall}{Precision+Recall} \dots\dots\dots(3)$$

Where:

Precision (Positive Predictive Value)

Recall (Sensitivity)

TP = True Positives (correctly predicted threats)

FP = False Positives (incorrectly predicted threats)

5.4 Recall:

Recall (also known as Sensitivity or True Positive Rate (TPR)) is a key evaluation metric in classification, especially for threat detection systems where identifying actual threats is crucial.

Recall Formula

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots(4)$$

Where:

1. TP (True Positives): Correctly predicted positive cases (actual threats detected correctly).
2. FN (False Negatives): Missed positive cases (actual threats incorrectly classified as normal traffic).

**Comparative Analysis:**

Model	Data Set	Accuracy	Precision	Recall	F1 Score
DNN(multi-class classification) [8]	NSL-KDD	77.80	78.00	77.8	76
CNN[10]	UNSW-NB15	82.48%	82.40%	82.48%	80.04%
DNN-1[17]	KDDCup-99	92.90	99.80	91.5	95.4
TFT(Proposed)	Simargyl	95.40	91.01	95.40	93.15

Table 1: Comparative Analysis

Comparative Analysis:

Intrusion detection systems (IDS) play a crucial role in cybersecurity by identifying malicious network activities. Deep learning techniques such as Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Temporal Fusion Transformers (TFT) have been widely used for network breach detection. This study aims to compare these models across different datasets to evaluate their effectiveness in detecting intrusions.

The TFT-based model outperforms all other models in terms of accuracy (95.40%), recall (95.40%), and F1-score (93.15%), making it the most effective approach for intrusion detection. While DNN-1 model on KDDCup-99, while achieving the highest precision (99.80%), suffers from a lower recall (91.5%), meaning it may fail to detect certain attacks, making it less reliable for comprehensive intrusion detection. CNN on UNSW-NB15, with an accuracy of 82.48% and a lower F1-score (80.04%), struggles to balance precision and recall, likely due to its reliance on spatial feature extraction, which is less effective for sequential network traffic patterns.

The choice of dataset significantly affects the performance of the models. KDDCup-99 and Simargyl allow higher accuracy, likely due to their well-structured feature sets and data distributions, which provide clear distinctions between normal and attack traffic.

The proposed TFT model excels in capturing long-term dependencies and behavioral patterns in network traffic, setting it apart from CNNs and DNNs, which primarily focus on static or spatial features. This temporal awareness enables TFT to recognize subtle attack patterns that evolve over time, improving its ability to detect both known and unknown threats.

The TFT-based model achieves state-of-the-art performance in network intrusion detection, outperforming DNN and CNN approaches. By leveraging temporal dependencies, feature attention, and a balanced precision-recall trade-off, TFT provides a more robust and efficient solution for cybersecurity applications. Additionally, the Simargyl dataset's realistic attack scenarios enhance the effectiveness of TFT, making it a superior choice for modern intrusion detection systems.

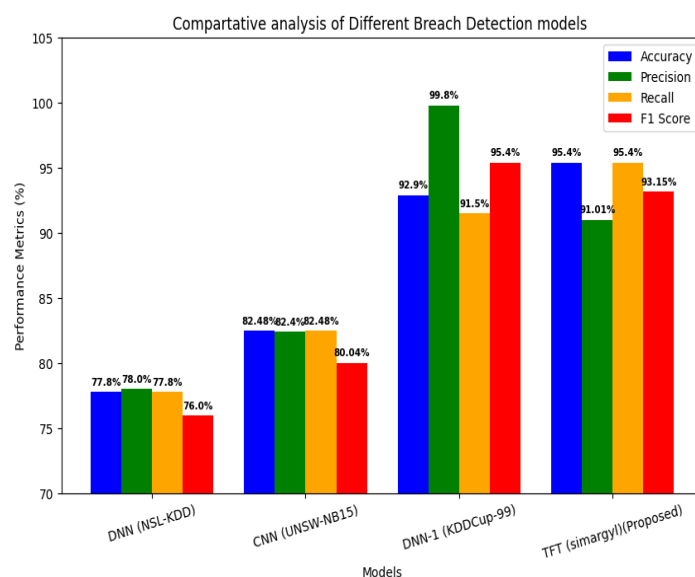


Fig-3: Bar Chart For Comparative Analysis



CONCLUSION

Research on using Temporal Fusion Transformers (TFT) for Network Intrusion Detection Systems (NIDS) is still in its early stages. TFT's ability to capture long-term dependencies and temporal patterns in sequential network traffic makes it a promising candidate for enhancing intrusion detection systems. By leveraging attention mechanisms and multi-horizon forecasting, TFT can effectively differentiate between normal and malicious network behaviour, improving detection accuracy while minimizing false positives. This study demonstrates that integrating TFT into NIDS can enhance cybersecurity defences, offering a robust, adaptive, and scalable approach to breach detection.

The experimental results of this research demonstrate that the proposed TFT-based breach detection system achieves high detection accuracy while reducing false alarms. Compared to traditional IDSs and machine learning-based methods, the TFT model provides superior performance in detecting network intrusions. This research highlights the importance of integrating deep learning and time-series analysis into modern cybersecurity systems to create adaptive, reliable, and efficient breach detection solutions. By leveraging advanced AI techniques, this study aims to enhance network security and provide organizations with a more effective approach to detect and prevent cyberthreats.

Future research should focus on optimizing TFT for real-time deployment in large-scale networks, improving its interpretability, and exploring hybrid models to strengthen network security against evolving cyber threats.

ACKNOWLEDGEMENT

We are deeply grateful to Vasireddy Venkatadri Institute of Technology, Nambur, Guntur for their unwavering support and guidance throughout this research project. Their expertise and guidance have significantly contributed to the development and enhancement of our Temporal Fusion Transformer (TFT) based breach detection model. We also extend our gratitude to Vasireddy Venkatadri Institute of Technology for providing the necessary resources and datasets crucial for training, testing, and evaluating our intrusion detection system. Their support and collaboration have been instrumental in the success of this research. Additionally, we acknowledge the valuable feedback, encouragement, and assistance from our colleagues and peers, which have played a vital role in refining our work. Their valuable insights and continuous support have played a crucial role in shaping the progress and results of our research. Additionally, we appreciate the contributions of the research community in driving advancements in cybersecurity and machine learning. The collective dedication to innovation and knowledge-sharing serves as a constant source of inspiration for our work.

REFERENCES

- [1] Hnamte, V., & Hussain, J. (2023). Dependable intrusion detection system using deep convolutional neural network: A novel framework and performance evaluation approach. *Telematics and Informatics Reports*, 11, 100077..
- [2] Mohammadpour, L., Ling, T. C., Liew, C. S., & Aryanfar, A. (2022). A survey of CNN-based network intrusion detection. *Applied Sciences*, 12(16), 8162. .
- [3] Sun, P., Liu, P., Li, Q., Liu, C., Lu, X., Hao, R., & Chen, J. (2020). DL-IDS: Extracting features using CNN- LSTM hybrid network for intrusion detection system. *Security and communication networks*, 2020, 1-11.
- [4] AYUSH CHOUBEY, Addapalli and VN Krishna. "Intrusion detection system using deep learning methodologies." *Journal of Mathematical and Computational Science* (2021): n. pag.
- [5] Altunay, H. C., & Albayrak, Z. (2023). A hybrid CNN+ LSTM-based intrusion detection system for industrial IoT networks. *Engineering Science and Technology, an International Journal*, 38, 101322.
- [6] Ahmad, Zeeshan & Shahid Khan, Adnan & Shiang, Cheah & Ahmad, Farhan. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies*. 32. 10.1002/ett.4150.
- [7] Ashiku, L., & Dagli, C. (2021). Network intrusion detection system using deep learning. *Procedia Computer Science*, 185, 239-247.
- [8] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," in *IEEE Access*, vol. 7, pp. 41525-41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
- [9] Laghrissi, Fatimaezzahra & Douzi, Samira & Khadija, Douzi & Hssina, Badr. (2021). Intrusion detection systems using long short-term memory (LSTM). *Journal of Big Data*. 8. 10.1186/s40537- 021-00448-4.
- [10] Jayalaxmi, P. L. S., Saha, R., Kumar, G., Conti, M., & Kim, T. H. (2022). Machine and deep learning solutions for intrusion detection and prevention in IoTs: A survey. *IEEE Access*, 10, 121173- 121192.
- [11] Banaamah, A. M., & Ahmad, I. (2022). Intrusion detection in iot using deep learning. *Sensors*, 22(21), 8417.



- [12] Hnamte, V., & Hussain, J. (2023). DCNNBiLSTM: An efficient hybrid deep learning-based intrusion detection system. *Telematics and Informatics Reports*, 10, 100053.
- [13] Ayantayo, A., Kaur, A., Kour, A., Schmoor, X., Shah, F., Vickers, I., ... & Abdelsamea, M. M. (2023). Network intrusion detection using feature fusion with deep learning. *Journal of Big Data*, 10(1), 167.
- [14] M. Verkerken, J. Santos, L. D'Hooge, T. Wauters, B. Volckaert and F. De Turck, "ChronosGuard: A Hierarchical Machine Learning Intrusion Detection System for Modern Clouds," 2024 20th International Conference on Network and Service Management (CNSM), Prague, Czech Republic, 2024
- [15] Y. V. Narayana and M. Sreedevi, "DDCATF: Deep Learning Approach for Detection of Cybercrime Activities Based on Temporal Features," 2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), Erode, India, 2023, pp. 462-469, doi: 10.1109/ICSSAS57918.2023.10331644.
- [16] R. K. Vigneswaran, R. Vinayakumar, K. P. Soman and P. Poornachandran, "Evaluating Shallow and Deep Neural Networks for Network Intrusion Detection Systems in Cyber Security," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 2018, pp. 1-6, doi: 10.1109/ICCCNT.2018.8494096.
- [17] Y. Zhang, H. Chen, and X. Liu, "Deep Learning-Based Anomaly Detection for Network Security," *Journal of Cybersecurity and Privacy*, vol. 2, no. 1, pp. 35-50, 2020.
- [18] X. Huang, L. Zhou, J. Zhang, and H. Wang, "Transformer-Based Models for Cyber Threat Intelligence and Anomaly Detection," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 1023-1035, 2022.
- [19] Lim, Bryan, et al. "Temporal fusion transformers for interpretable multi-horizon time series forecasting." *International Journal of Forecasting* 37.4 (2021): 1748-1764.