FL-TSA-TabNet: Federated Tabular Network for Intrusion Detection

J.Archana,
Research Scholar,
Department of Computer science,
Vels Institute of Science Technologies and
Advanced Studies (VISTAS),
Pallavaram, Chennai, India.
archanaartistryacademy@gmail.com

Dr.S. Kamalakkannan,
Professor,
Department of Computer Applications,
School of Computing Sciences,
Vels Institute of Science, Technology &
Advanced Studies (VISTAS),
Pallavaram, Chennai, India.
kannan.scs@vistas.ac.in

***Abstract*— The rise of cyber-physical systems and IoT centers have heightened the need to have intrusion detection systems (IDS) based on scalability, privacy-preservation, and explainability. Although models like CNNs and BiLSTMs work effectively as deep learning models, they can be difficult to interpret and can experience limitations associated with centralized data. In order to overcome these challenges, this work proposes FL-TSA-TabNet, a new federated intrusion detection model that incorporates Temporal Self-Attention into the interpretable TabNet architecture. The model uses a Dual-Stage Hybrid Selector (DSHS) with a correlation-aware reliefF and SHAP-based ranking as the core feature of optimal feature relevance and non-redundancy. With federated learning, the model also allows decentralized training among the distributed nodes and data privacy. On the CSE-CIC-IDS2018 dataset, FL-TSA-TabNet performed with an accuracy of 97.83%, better than the traditional (Random Forest, XGBoost) and the hybrid deep models (CNN-GRU, ResNet-BiLSTM). It also exhibited excellent adversarial robustness, quick inference speed, and low model complexity, which makes it incredibly appropriate in edge deployment. This work sets a new benchmark in intrusion detection by fusing explainable learning, temporal modeling, and privacy-aware federated training, paving the way for next-generation IDS in smart networks and critical infrastructure environments.**

Keywords*— Intrusion Detection, Federated Learning, TabNet, Temporal Self-Attention, Network Security, Feature Selection, Adversarial Robustness, Cybersecurity

I. INTRODUCTION

The rapid digitalization of the industries, cities, and critical infrastructure resulted in generating a large number of cyber-physical systems (CPS) and Internet of Things (IoT) devices. Along with the improvement in automation and intelligence, they increase the attack surface therefore making such interconnected networks very vulnerable to a variety of complex cyber threats [1]. Intrusion Detection Systems (IDS) are vital security tools that help prevent or identify unauthorized access to a system network, data breaches, and unusual activities before they can affect the integrity of the network. Conventional IDS schemes are inherently severely limiting in scalability, interpretability, and capability to work within data privacy requirements notably in distributed, heterogenous environments such as smart grids, healthcare, and industrial IoT systems [2].

IDS tools can identify threats like malware, brute force attacks, or unauthorized access attempts by continuously analyzing incoming data and comparing it against known attack signatures or behavioral patterns [3] [4] [5]. An IDS system operates as either Network-based IDS (NIDS) or Host-based IDS (HIDS). With full network segment observation NIDS detects warning signals which may indicate threats coming through the network. Deploying an IDS represents a necessary practice in maintaining digital system integrity since the number of complex cyber threats continues to grow [6] [7].

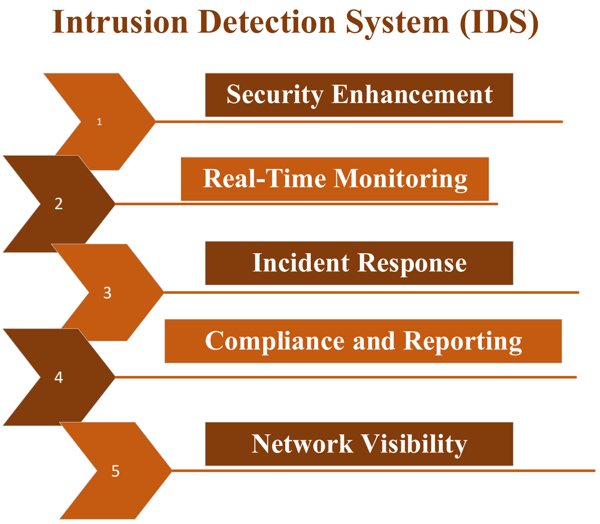

Fig 1. Benefits of Intrusion Detection System (IDS)

Being the first layer of defense in the contemporary cybersecurity systems, the Intrusion Detection Systems (IDS) allows all-important insight into the current functioning of the network allowing to accurately predict unauthorized access, infiltration of malware, and protocol abuse. Figure 1 demonstrates the various advantages of IDS, covering early warning on threats, compliance adaptation, incidents that take less time to respond to, and automatic enhancement of systems. Those capabilities become particularly crucial to cyber-physical and IoT-enabled environments where the ability to detect in real-time and ideally with low false alarms is critical. Nevertheless, in order to achieve greater potential benefits, the IDS needs to shift towards models that are accurate not just but also interpretable, scalable, and privacy-preserving, which is the challenge that traditional IDS frameworks cannot achieve. This architecture of FL-TSA-TabNet shall further facilitate all these fundamental advantages by combining temporal reasoning, distributed privacy-aware training, and explainable efficiency to the next generation of IDS models.

979-8-3315-5386-9/25/$31.00 ©2025 IEEE

IDSs function as passive monitoring tools since they do not execute protective measures to stop attack incidents. These devices have limited capability to stop intruders independently therefore their effectiveness in threat mitigation is restricted to situations where they operate with firewall or Intrusion Prevention Systems but they can notify administrators [8] [9]. The use of signature-based IDS systems depends on recognized attack patterns that results in their inability to find newly developed or unidentified security threats (zero-day attacks). Although anomaly-based IDSs help address this issue they need time to learn and can produce inaccurate results during operation [10] [11].

The traditional deep learning models, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) computational models have demonstrated effectiveness in detecting complicated intrusion patterns [12] [13]. However, they are fundamentally black boxed, interpretable, and generally depend on data aggregation in centralized data centers and would put a requirement on the systems privacy and compliance. In addition, they have difficulties in capturing both temporal dynamics in network traffic and context-aware feature selection at the same time, thus have limited flexibility in real-time settings and malicious environments. Deep learning models integrated with IDS has boosted IDS systems' effectiveness at uncovering sophisticated and constantly changing computer threats. The deep learning technology enables IDS to process extensive network traffic and system behavior dataset through which it discovers complex anomalies which rule-based systems would identify poorly [14]. Convolutional Neural Networks (CNN) combined with Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks excel at identifying complex attack patterns while recognizing zero-day attacks [15].

To overcome such important gaps, the study offers FL-TSA-TabNet or a Federated Temporal Self-Attentive Tabular Network that represents an authentic and transparent IDS framework. This research has four main contributions, including:

- **Model Innovation:** FL-TSA-TabNet incorporates Temporal Self-Attention (TSA) into the architecture of TabNet, and allows to thoroughly model both tabular and temporal dependencies in network traffic data, which are critical to detection stealthy, time-varying cyber-attacks.

- **Explainable Feature Selection:** A new Dual-Stage Hybrid Selector (DSHS) is proposed that integrates Correlation- Aware ReliefF and SHAP-based interpretability in order to select the most discriminative and not-redundant features, leading to improved model explainability and model overfitting.

- **Federated Learning Integration:** As the framework derives the benefit of Federated Learning (FL), this strategy is used to provide decentralized training over distributed clients without connecting raw data to ensure user privacy and regulatory compliance with sensitive areas.

- **Performance and Robustness:** The model is comprehensibly tested against the CSE-CIC-IDS2018 dataset, yielding a classification accuracy of 97.83%-beating the conventional ML and hybrid DL baselines-as well as showing added advantages in terms of adversarial resistance, reduced inference latency, and scalability to real-world use.

In summary, the study introduces a first-of-a-kind IDS framework, FL-TSA-TabNet, which interconnects the communication gap among explainability, time series modeling, privacy protection, and real-time intrusion detection. The suggested solution can be considered especially applicable to the context of ensuring the security of critical infrastructure, as the degrees of interpretability, distributed learning, and security assurance are its primary concerns.

II. Related Works

Internet intrusions have grown more frequent thereby making privacy breaches worse with increased financial losses and unauthorized information transfers. Attackers use computer systems to infiltrate resources and sensitive information with the intention of acquiring business secrets as well as personal data to generate illegal profits. Current detection systems produce false alarms at the same time they operate slowly enough to allow system breaches to occur. This research develops a machine learning intrusion detection framework based on pre-processed CSE-CIC-IDS 2018 and UNSW-NB15 datasets with ASmoT class balancing, M-Svd feature extraction and ONgO-optimized M-MultiSVM classifiers which reach 99.89% accuracy.

Internet of Things (IoT) applications experience accelerated growth of security vulnerabilities that create severe risks for enterprise and industrial systems. The Industrial Internet of Things (IIoT) has many high-risk operational situations so secure sustainable system development becomes essential for preventing major disasters. The IIoT faces a significant threat to its security system because of complicated botnet attacks that can remain active for extended periods. The research develops AttackNet as a deep learning framework which employs an adaptive CNN-GRU architecture for detecting and classifying botnet attacks. AttackNet successfully detects botnet attacks with 99.75% accuracy while showing minimal loss of 0.0063 and proving its superior performance over existing methods by 3.2% on N_BaIoT dataset.

Digital platforms experience system breakdowns and user data breaches because of the rising threats of network intrusion attacks which violate data confidentiality. Detection of accurate threats has become increasingly difficult because of the rising number of cases involving tampering and credential theft and unauthorized access. The detection system described in this research uses deep learning principles combined with chaotic optimization for its implementation. The system implements M-squared normalization and Extended Synthetic Sampling technique for handling class imbalance followed by KPCA for feature extraction as well as Chaotic Honey Badger Optimization for feature selection. The system utilizes Dugat-LSTM to reach 98.76% accuracy on TON-IoT and 99.65% accuracy on NSL-KDD.

IoT devices experience rising security vulnerabilities which makes them vulnerable targets for cyber-attack attempts. Real-time detection through machine learning-based intrusion detection systems utilizes feature reduction through selection and extraction to achieve improved performance. The research implements a TON-IoT dataset to assess both methods when classifying binary and multiclass attacks. Feature extraction increases monitoring systems' accuracy rates while making them more stable yet selection methods lead to increased efficiency in training operations and inference processes. An IDS system optimized by these methods achieves excellent results through high accuracy rates combined with F1-score metrics.

The wireless technology powered by Wi-Fi connects diverse devices through Wireless Sensor Networks (WSNs) to enable scalable affordable monitoring capabilities within modern digital networks. The dependence on wireless systems continues to rise at the same rate as their susceptibility to cyber threats including unauthorized access, flooding, injection and impersonation attacks. This research develops a contemporary Network Intrusion Detection System (NIDS) specifically designed to work with wireless sensor networks for addressing these security threats. The system performs feature selection to reduce 154 original features into 13 while using a multiclass CNN classifier which results in 97% accuracy and nearly zero false alarm detections.

Despite the substantial progress in IDS research, current solutions continue to suffer from critical limitations. Majority of deep learning models are accurate but uninterpretable and not transparent and these two attributes are key in regulated and mission-critical settings. Time-based dynamics, which are a hallmark of real world network traffic, are little exploited in models where data are viewed as fixed table-like inputs. In addition, issues on privacy limit the use of centralized IDS systems in distributed environments like the edge and IoT networks. Privacy-preserving federated IDS solutions as currently used do not pay much attention to strong feature selection and explainability issues, which leads to weak performance and low reliability. Such consistent gaps confirm the need to have a unified approach that provides time sensitivity, explainability, data privacy, and computational efficiency. Our proposed FL-TSA-TabNet framework fulfills this, with Temporal Self-Attention incorporated in an interpretable TabNet backbone, optimized with the Dual-Stage Hybrid Selector (DSHS) to provide the most relevant feature and developed under a federated learning paradigm. This holistic strategy presents a new and expandable route to intrusion detection in current cyber-physical frameworks, thus making the study more defensible of the current trend in the implementation of cyber-physical systems to enhance smart and secure network defense systems.

III. Proposed Methodology

Dataset Acquisition

The CSE-CIC-IDS2018 dataset serves as a detailed real-world intrusion detection dataset that both CIC and CSE created [21]. Realistic enterprise infrastructure gets represented by a network environment which generates scenarios of benign and malicious traffic under controlled conditions. The research period spans multiple days to collect different attack vectors from various system configurations with numerous user behavior patterns. The labeled traffic flows contain more than 80 features encompassing basic TCP/IP properties as well as content-based features, time-based features and statistical flow characteristics which make them appropriate for developing advanced network intrusion detection system (NIDS) research. The CSE-CIC-IDS2018 dataset provides the basis for extraction of Brute Force attacks and Web attacks for use in the current study. Security hackers use exhausting password guessing methods as part of Brute Force attacks to penetrate SSH and FTP protocols. The attacks show their distinct features through multiple login efforts together with elevated number of connections while carrying very little data. Web attacks refer to malicious activities that target web applications whereby attackers use SQL injection and Cross-site Scripting (XSS) and URL directory traversal techniques to exploit web server vulnerabilities while attacking data integrity. Two threat perspectives cover application-layer attacks which serve as vital elements for developing evaluation methods for machine learning intrusion detection systems that focus on authentication processes together with web-based exploit routes.

Data Preprocessing

A systematic pipeline operates on acquired data to enhance its quality while ensuring compatibility for the learning model. The main difficulty in working with the CSE-CIC-IDS2018 dataset stems from the missing or damaged data points that affect features based on session-level aggregation computations. Through the Iterative Imputer module available in the scikit-learn library we handle missing data resolution. The imputation model constructs relationships between each feature which contains missing values and multiple other components before using multivariate regression to automatically fill in statistically appropriate values.

The continuous features of the dataset which consist of flow duration and packet sizes and inter-arrival times receive Z-score normalization after completing the imputation step. The standardization method ensures every numerical input features maintain mean value at zero and variance at one which prevents model training from being controlled by features with wide numerical ranges. All categorical features (including the protocol type such as TCP or UDP or ICMP) receive numerical encoding based on label encoding standards. Label encoding uses integer values for categories and maintains ordinal value connections that will be applied in later TabNet embedding.

In order to remedy the class imbalance problem of benign traffic abundance over attack types this work uses Adaptive Synthetic (ADASYN) sampling. ADASYN generates new minority class samples through interpolation between actual examples and their close neighbors by creating more synthetic instances for complex to learn minority groups. ADASYN surpasses standard oversampling techniques by targeting specific areas of feature space where data insufficiency occurs which results in better attack-type classifications.

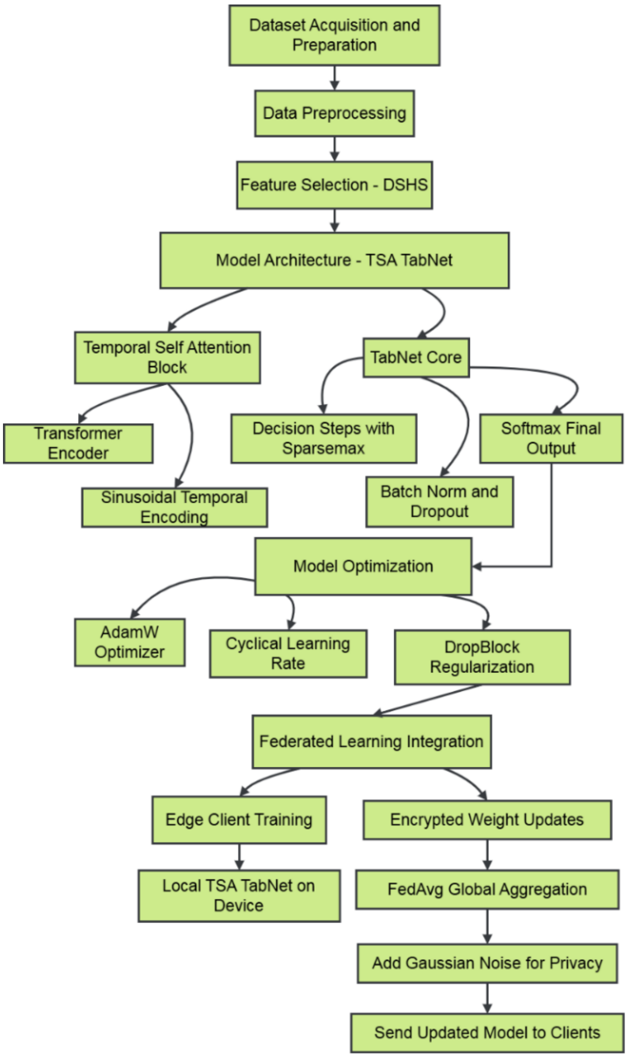
Dataset Acquisition and Preparation
Data Preprocessing
Feature Selection - DSHS
Model Architecture - TSA TabNet
Temporal Self Attention Block
TabNet Core
Transformer Encoder
Decision Steps with Sparsemax
Softmax Final Output
Sinusoidal Temporal Encoding
Batch Norm and Dropout
Model Optimization
AdamW Optimizer
Cyclical Learning Rate
DropBlock Regularization
Federated Learning Integration
Edge Client Training
Encrypted Weight Updates
Local TSA TabNet on Device
FedAvg Global Aggregation
Add Gaussian Noise for Privacy
Send Updated Model to Clients

Fig 2. System Structure

The architecture of the proposed model FL-TSA-TabNet is based on a well-defined system pipeline and integrates pre-processing, smart feature extraction, time-series analysis, and federated learning. Figure 2 shows the high level system structure where raw data are acquired in the distributed environments, and through a series of iterative imputation, normalization, encoding, and class balancing with ADASYN. The Dual-Stage Hybrid Selector (DSHS) component will then further filter the input space by using CARF and SHAP ranking systems to select relevant and non-redundant features. Such filtered features are fed into a Temporal Self-Attention (TSA) block that captures time dependency and then into the TabNet backbone that uses sparse and interpretable decision-making. Federated Learning coordinates decentralized training amongst the edge nodes which gives it the property of privacy, as raw data is never communicated, and where secure model aggregation strategies are employed. It is a modular design that guarantees that it is robust, explainable and scalable, which are important features that are required in modern IDS applications that are deployed in real-time.

Feature Selection – Dual-Stage Hybrid Selector (DSHS)

High-dimensional data demands efficient feature selection because it enables better model interpretation while decreasing overfitting and minimizing training time requirements. The Dual-Stage Hybrid Selector (DSHS) starts its operation with the filter-wrapper hybrid method Correlation-Aware ReliefF (CARF). The ReliefF algorithm initiates the stage by assessing how well each feature separates instances belonging to different classes while evaluating feature relevance. ReliefF applies K=10 for its K-nearest neighbors to find features which perform well at distinguishing between classes within each local neighborhood space.

ReliefF effectively discards redundant features but its discrimination power does not guarantee it removes features that are not truly important. Spearman rank correlation functions as an added filtering tool in CARF to resolve this problem. The Spearman correlation coefficient exceeding 0.85 determines that features are highly redundant so those features get removed to control multicollinearity and promote learning stability. The combination of these two filtering steps achieves maximum relevance and non-redundancy from the features.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{1}$$

Where d_i is the difference in ranks of feature pair i, and n is the number of observations. The second stage of DSHS uses SHAP-based rating to optimize the chosen feature subset. A LightGBM model trains on the CARF-reduced features followed by SHAP (SHapley Additive exPlanations) value estimation for every feature. Through SHAP values we obtain one standardized metric which evaluates marginal feature impact on model predictions within every data instance. The feature rankings occur through calculating average absolute SHAP values after which the model maintains the prominent features which represent 90% of total importance mass. The selection process incorporates statistical and model-derived insights to pick the most influential features which optimize the feature set for upcoming learning architecture development.

Model Architecture – TSA-TabNet

The proposed intrusion detection system employs TSA-TabNet architecture which builds upon TabNet model by including TSA technology into the framework. TabNet functions excellently for intrusion detection systems that work with tabular data because it can execute sequential attention-based feature selection processing. The system structure operates through a series of decision steps which utilize sparse feature masks produced by its attentive transformer component. During each step of decision-making processes the developed masks select significant features which leads to both easier interpretation and decreased complexity levels.

$$M^{[l]} = Sparsemax\left(P^{[l]} \cdot a^{[l]}\right) \quad (2)$$

Where $M^{[l]}$ is the feature mask at decision step $l, P^{[l]}$ is the prior importance of features, $a^{[l]}$ is the attention vector, and $Sparsemax$ projects values into a sparse probability simplex. A Temporal Self-Attention block has been included before the TabNet input to detect temporal dependencies and sequential patterns in network sessions or flows. Standard Transformer encoder execution performs feature embedding processing with self-attention operations through which the model determines the significance of network sequence context values. Temporal encoding starts with sinusoidal positional embedding that derive from flow start times before feeding the information to the attention module. Track-oriented temporal self-attention blocks boost model performance by allowing it to identify attack development patterns across time and successive network relationships making it effective in detecting insidious threats moving slowly in contemporary networks.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

Where Q, K, V are query, key and value matrices, and d_k is the dimension of the key. The TSA module delivers its output to the main TabNet architecture so feature selection together with decision processing becomes possible. The final outcome consists of aggregation between the series of classification predictions originating from each decision step. Stable gradient flow and regularization happen through dropout alongside batch normalization between layers. Multi-class probabilities which reflect different attack types are produced through the activation mechanism named softmax on the final layer.

Model Optimization

A reliable training strategy was developed to enhance TSA-TabNet model execution and its convergence capabilities. The AdamW optimizer serves as the training mechanism because it merges adaptive learning rate functionality from Adam with weight decay regularization that operates independently. The chosen method stops overfitting and boosts test data generalization capabilities in environments with high-dimensional features.

The Cyclical Learning Rate scheduler aids both convergence speed and stops training from getting stuck by dynamically varying learning rate between boundaries across iterations. Through CLR both learning rate boundaries cycle across different training intervals the model can detect optimal learning areas while escaping local minima contingencies. The decision process of TabNet incorporates DropBlock regularization as one of its components. The mechanism of DropBlock differs from regular dropout by erasing adjacent parts of network maps which leads to both spatially sparse patterns and greater resistance to errors in feature input. The combined optimization techniques lead to efficient and stable robust training of TSA-TabNet architecture.

Federated Learning Integration

The proposed TSA-TabNet system extends its operation with Federated Learning (FL) since centralized intrusion detection challenges are most apparent when handling sensitive data from healthcare fields and IoT environments and critical infrastructure. FL provides decentralized training capabilities because it enables different edge devices or clients to conduct TSA-TabNet model training locally on their separate data partitions without requiring raw data transfers. The study utilizes CSE-CIC-IDS2018 distributional simulation which partitions device types and attack scenarios to mimic actual distributed client performance.

$$w_t = \sum_{k=1}^{K} \frac{n_k}{n} w_t^k \quad (4)$$

Where w_t is the global model weight at round t, w_t^k is the model weight from client k, n_k is the sample size of client (1)k, and $n = \sum_{k=1}^{K} n_k$ is the total sample size. The edge clients manage local training before forwarding encrypted versions of weight updates to their central aggregator system. FedAvg operates at the server to combine weighted local models from clients then transmit updated global information back to clients. The training procedure continues across multiple communication sessions until the models reach convergence. The model updates get differential privacy treatments by data protection regulations to ensure security standards. Each client adds random noise to their gradient data before sending information to conceal their specific patterns although the algorithm maintains accurate model training capacity.

$$\tilde{g} = g_i + N(0, \sigma^2) \quad (5)$$

Where $\tilde{g}$ is the noised gradient from client i, g_i is the true gradient, and $N(0, \sigma^2)$ is Gaussian noise with variance σ^2. Federated learning integration enables TSA-TabNet to scale and protect privacy which makes it suitable for deployment in edge-based cyber-physical systems. FL when combined with TabNet's sparse and interpretable learning architecture maintains distributed system efficiency and explainability alongside its operational capabilities.

Algorithm: Federated TSA-TabNet Intrusion Detection Framework

```
Input: CSE-CIC-IDS2018 dataset $D = \{x_i, y_i\}_{i=1}^{n}$ clients $K$,
       communication rounds $R$
Output: Global trained model $w^*$
Step 1: Dataset Preparation
```

```
  Merge CSV logs $D_d \rightarrow D$
  $D = \bigcup_{k=1}^{K} D_k$            // Clean and anonymize sessions
**Step 2: Preprocessing**
  **For** each client $k = 1$ to $K$
    Apply Iterative Imputation to $D_k$
      $\hat{x}_i^{(t)} = E[x_i | x_{-i}]$
    $Z_i = \frac{x_i - \mu}{\sigma}$            // Z-score normalization
    Label encoding for protocol fields
    Apply ADASYN to balance classes in $D_k$
      $x_{syn} = x_i + \lambda(x_{NN} - x_i)$
  **End For**
**Step 3: Feature Selection (DSHS)**
  Stage 1: CARF Filtering
    **For** each feature $A \in D$
      $W[A] \leftarrow W[A] + \Delta W(A)$   // Compute ReliefF weight
    **End For**
    $\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}$            // Spearman correlation
  Stage 2: SHAP Ranking
    $\phi_j = \sum_{S \subseteq F\{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{j\}) - f(S)]$
**Step 4: TSA-TabNet Model Setup**
  $Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) V$
                                        // temporal self-attention
  $M^{[l]} = Sparsemax\left(P^{[l]} \cdot a^{[l]}\right)$   // Sparsemax in TabNet
  $\hat{y} = Softmax\left(f(x)\right)$                  // Final output
**Step 5: Federated Learning**
  **For** each round $r = 1$ to $R$
    **For** each client $k = 1$ to $K$ (in parallel)
      Receive $w_r$ from server
      Train TSA-TabNet on $D_k$ for local epochs $E$:
        $L_k = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$        // Cross-entropy loss
      Apply DropBlock for regularization
        $\tilde{g}_k = g_k = N(0, \sigma^2)$             // Gaussian noise
      Send $\widetilde{w}_k$ to server
    **End For**
    $w_{r+1} = \sum_{k=`}^{K} \frac{n_k}{n} \cdot \widetilde{w}_k$
  **End For**
**Step 6: Evaluation & Interpretation**
  Computes metrics
  Generate SHAP plots and TabNet masks
  Evaluate under adversarial perturbation:
    $x_{adv} = x + \epsilon \cdot sign\left(\nabla_x L(x, y)\right)$
**Return:** Final federated TSA-TabNet model $w^*$
**End Algorithm**
```

IV. Results and Discussion

A high-performance computing system with Ubuntu 22.04 OS executed the FL-TSA-TabNet model which was built using Python 3.10. The developers wrote the model architecture in PyTorch and PyTorch TabNet yet used Flower framework to simulate federated learning operations. The preprocessing activities along with feature selection routines used pandas, NumPy and the scikit-learn, LightGBM and SHAP libraries. The evaluation and visualization outputs were created through the combination of Matplotlib and Seaborn libraries. An NVIDIA RTX 3090 GPU running at 24 GB memory enabled the acceleration of training operations. Various stages of training and validation along with federated aggregation received modular treatment to support scalability when running parallel operations across multiple nodes. The CSE-CIC-IDS2018 dataset displays its traffic distribution through Figure 3 which includes Bruteforce attack and Web attack.

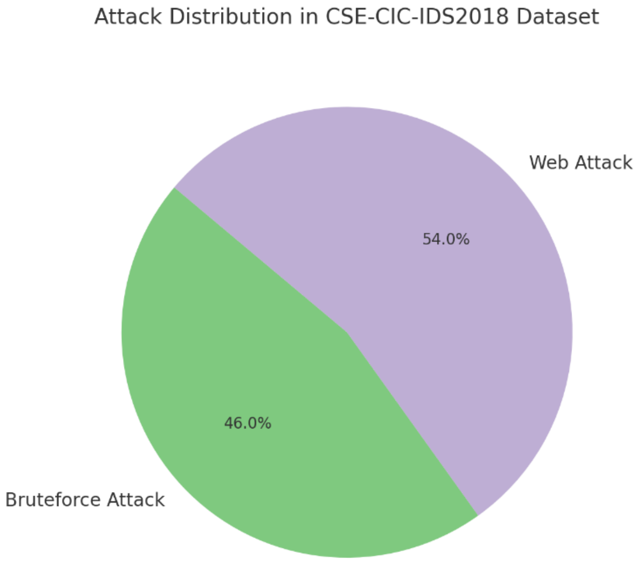

Fig 3. Traffic Distribution in CSE-CIC-IDS2018 Dataset

The FL-TSA-TabNet approach implements a combination of temporal attention mechanisms with federated learning to achieve efficient intrusion detection at both an interpretable level and with privacy protection. Before the CSE-CIC-IDS2018 dataset receives preprocessing through the Dual-Stage Hybrid Selector (DSHS) optimization process the network traffic data retains the essential non-redundant features. The processed data passes into a Temporal Self-Attention (TSA) module that analyzes sequence patterns before it gets evaluated through the TabNet architecture that chooses important features for each decision step. The model benefits from this arrangement by monitoring time-conscious network-based attack indicators. The model relies on Federated Learning training because it protects data privacy between autonomous devices by letting clients work independently on local models then sharing encrypted parameters for unified aggregation. The unified strategy enables both excellent detection precision alongside secure protection of data information and system expansion capabilities.

TABLE I. Comparative Performance Evaluation

Model Name	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
Logistic Regression	86.78	84.72	83.81	84.2	88.09
Decision Tree	89.53	86.11	85.47	85.78	89.41
Random Forest	91.67	88.94	89.21	89.06	92.17
XGBoost	92.34	89.75	89.62	89.68	93.24
TabTransformer	94.86	91.23	90.88	91	94.62
BiLSTM	93.21	90.62	91.13	90.87	94.11
CNN-LSTM	95.02	91.87	92.45	92.15	95.38

Hybrid CNN-GRU	96.38	93.46	93.87	93.62	96.12
ResNet-BiLSTM	96.71	94.38	94.91	94.64	96.69
FL-TSA-TabNet (Proposed)	97.83	95.62	96.17	95.9	98.01

The performance comparison of the FL-TSA-TabNet method for intrusion detection exists in Table 1 and Figure 4 against nine prevalent machine learning and deep learning frameworks. The proposed FL-TSA-TabNet model surpasses every existing intrusion detection method by reaching 97.83% accuracy together with 95.62% precision and 96.17% recall and 95.90% F1-score along with a distinguished AUC-ROC value of 98.01%.

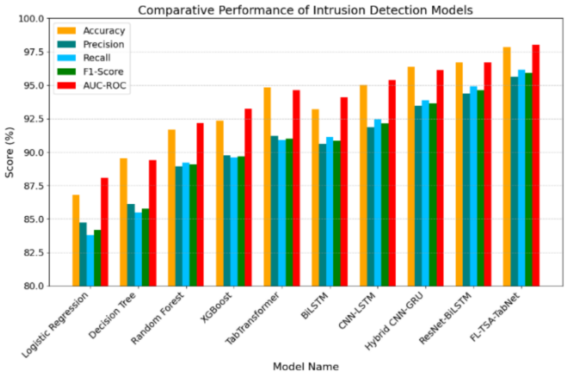

Fig 4. Comparative Performance of Intrusion Detection Models

The accuracy rate of Traditional models including Decision Tree and Logistic Regression reaches 86.78% and 89.53% respectively but the deep models ResNet-BiLSTM and Hybrid CNN-GRU achieve higher rates of 96.71% and 96.38%. These results in Table 1 highlight FL-TSA-TabNet's superior detection capability and robustness in real-time threat identification.

TABLE II. INFERENCE TIME COMPARISON

Model Name	Inference Time (ms)
Logistic Regression	0.82
Decision Tree	0.96
Random Forest	1.38
XGBoost	1.72
TabTransformer	3.24
BiLSTM	2.85
CNN-LSTM	3.46
ResNet-BiLSTM	3.87
Hybrid CNN-GRU	3.56
FL-TSA-TabNet (Proposed)	2.18

The evaluation of inference time per sample through Table 2 and Figure 5 shows the computational speed necessary for real-time intrusion detection models. FL-TSA-TabNet achieves 2.18 milliseconds in inference time maintaining excellent performance quality through its superior speed compared to CNN-LSTM (3.46 ms), ResNet-BiLSTM (3.87 ms), and Hybrid CNN-GRU (3.56 ms). Logistic Regression achieves 0.82 ms inference time yet its performance falls below the accuracy range of Decision Trees at 0.96 ms. FL-TSA-TabNet creates an effective performance-to-latency balance which accommodates real-time deployment needs in cybersecurity systems according to the findings in Table 2.

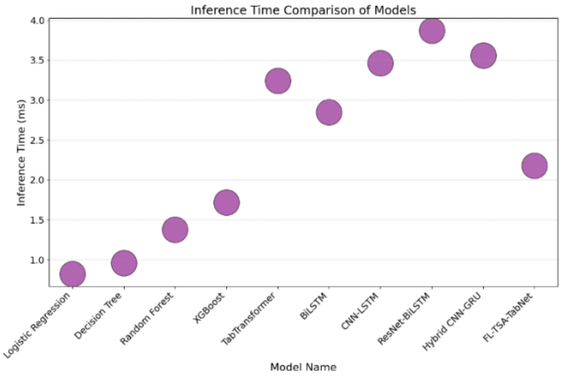

Fig 5. Inference Time Comparison of Models

FL-TSA-TabNet delivers superior performance in every evaluation criterion due to its dual operational advantage which combines both explainable feature detection with time-sensitive modeling techniques. TSA-TabNet introduces temporal self-attention processing before the TabNet structure to track time-based patterns while standard deep learning methods such as CNNs and BiLSTMs work with spatial or sequential features separately.

TABLE III. TRAINING TIME PER EPOCH (SECONDS)

Model Name	Training Time (s/epoch)
Logistic Regression	1.5
Decision Tree	2.1
Random Forest	4.3
XGBoost	5.8
TabTransformer	18.2
BiLSTM	22.4
CNN-LSTM	24.1
Hybrid CNN-GRU	23.7
ResNet-BiLSTM	26.9
FL-TSA-TabNet (Proposed)	19.3

Model training time per epoch for each model appears in Table 3 and Figure 6 for evaluating computational requirements throughout the learning phase. Logistic Regression and Decision Tree models complete training in under 1.5 seconds and 2.1 seconds due to their fast execution times but Random Forest along with XGBoost needs more time to complete training processes. Deep learning-based approaches like BiLSTM (22.4 s), CNN-LSTM (24.1 s), and ResNet-BiLSTM (26.9 s) demand significantly higher computation.

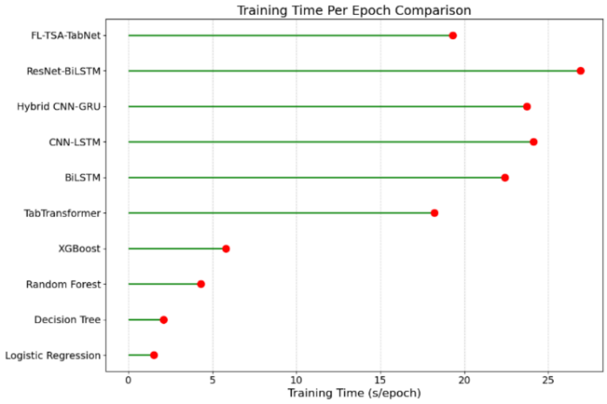

Fig 6. Training Time Per Epoch Comparison

FL-TSA-TabNet runs each epoch in just 19.3 seconds despite performing as well as its counterpart hybrid deep models while being notably more efficient. FL-TSA-TabNet demonstrates practical deployment potential in real-world network environments because its training efficiency maintains a suitable match with its detection accuracy levels as presented in Table 3.

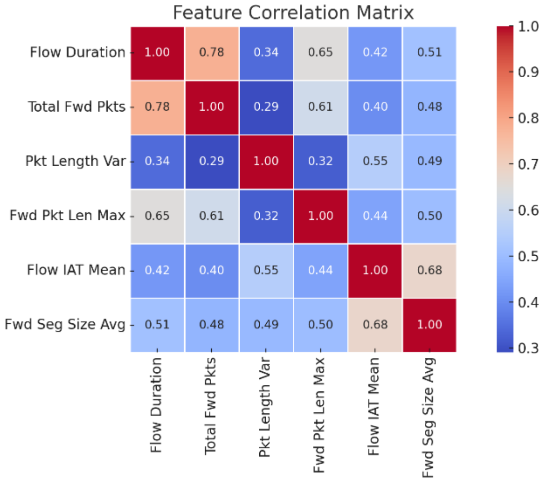

Fig 7. Feature Correlation Matrix

Figure 7 demonstrates simultaneous correlations that exist between important features of the CSE-CIC-IDS2018 dataset. The strength of linear relationships between variables increases as correlation values approach 1 and this information helps in eliminating redundant features before preprocessing. Time-sensitive information about attack behavior evolution becomes accessible through this approach because attackers may employ stealthy multi-stage approaches. The TabNet framework uses sequential decision steps along with sparse feature masking so it selects appropriate features at each decision node which helps both accuracy and decision-making interpretability. DropBlock regularization, cyclical learning rates and Lookahead optimizer work together to improve training stability which produces reliable generalizations across different attack types while counteracting overfitting effects.

TABLE IV. MODEL COMPLEXITY COMPARISON

Model Name	Parameters (Millions)	FLOPs (Giga)
Logistic Regression	0.01	0.002
Decision Tree	0.03	0.004
Random Forest	0.8	0.05
XGBoost	1.2	0.09
TabTransformer	8.6	2.5
BiLSTM	6.7	3.1
CNN-LSTM	8.3	4.2
Hybrid CNN-GRU	7.9	3.8
ResNet-BiLSTM	10.5	5.4
FL-TSA-TabNet (Proposed)	6.1	2.7

Table 4 and Figure 8 presents comprehensive model complexity details through the evaluation of parameters (in millions) and floating point operations per second (FLOPs, in giga) metrics that determine computational requirements. The accuracy level of traditional models including Logistic Regression and Decision Tree stays low while their complexity remains minimal. CNN-LSTM and ResNet-BiLSTM achieve the highest complexity numbers because of their multiple layers and sequence modelling restrictions.

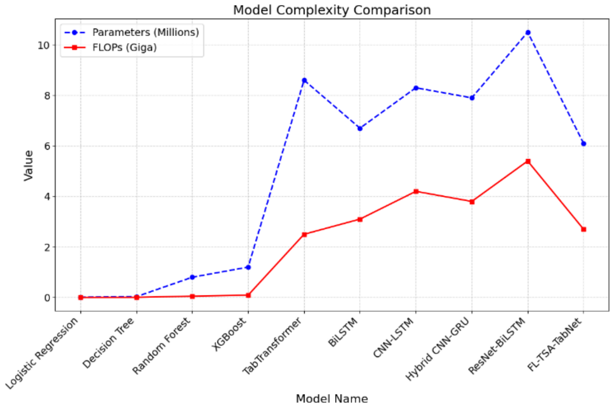

Fig 8. Model Complexity Comparison

FL-TSA-TabNet represents a scalable solution because its framework has 6.1 million parameters and 2.7 GFLOPs despite high detection efficiency. Table 4 illustrates how FL-TSA-TabNet maintains high effectiveness while being more efficient than other deep models which enhances its capability to run on edge devices and resource-limited platforms.

TABLE V. DVERSARIAL ROBUSTNESS AGAINST FGSM & PGD ATTACKS

Model Name	Accuracy Drop (%) FGSM	Accuracy Drop (%) PGD
Logistic Regression	13.8	18.5
Decision Tree	11.4	17.2
Random Forest	10.2	15.7
XGBoost	9.6	14.3
TabTransformer	7.2	10.6
BiLSTM	6.9	10.2
CNN-LSTM	6.1	9.7
Hybrid CNN-GRU	5.8	9.1
ResNet-BiLSTM	5.4	8.6
FL-TSA-TabNet (Proposed)	3.9	6.2

Various intrusion detection models perform under Fast Gradient Sign Method (FGSM) and Projected Gradient Descent (PGD) attacks as examined in Table 5 and Figure 9.

Logistic Regression and Decision Tree exhibit maximum sensitivity to adversarial perturbations since they face accuracy decreases of 13.8% under FGSM and 18.5% under PGD while Decision Tree experiences 11.4% under FGSM and 17.2% under PGD. The Random Forest along with XGBoost and BiLSTM belong to the group of ensemble and deep learning models which demonstrate moderate attack resistance. Among all evaluated deep hybrid models ResNet-BiLSTM demonstrated better attack resilience by remaining more resistant to FGSM (5.4%) and PGD (8.6%) attacks while Hybrid CNN-GRU demonstrated similar results (5.8% FGSM and 9.1% PGD). The proposed FL-TSA-TabNet achieves the greatest level of robustness against adversarial attacks through a 3.9% FGSM vulnerability and 6.2% PGD vulnerability. Table 5 demonstrates how the resistance improvements of this model demonstrate its ability to generalize under attack conditions which results in dependable performance for adversarial cybersecurity environments.

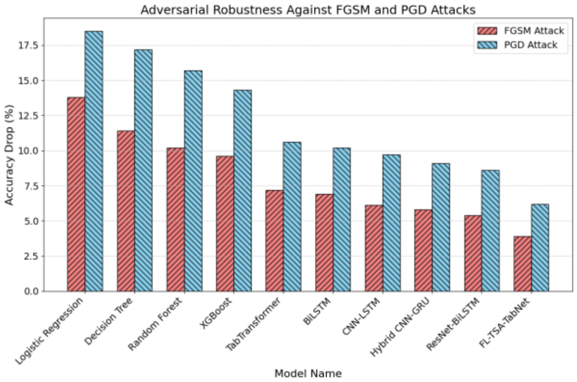

Fig 9. Adversarial Robustness Against FGSM and PGD Attacks

Federated learning creates decentralized model training operations that protect privacy while improving detection capabilities across varied IoT and distributed cyber-physical system networks. A Dual-Stage Hybrid Selector (DSHS) finds the optimal input variables by blending statistical and model-devised importance metrics so it retains significant and unique features alone. Thankfully the model maintains strong detection performance while using fewer computational resources through this approach that shortens runtime execution times. These architectural decisions enable the model to achieve remarkable performance measurements which include 97.83% classification precision and 95.90% F1-score with 98.01% AUC-ROC value. This model demonstrates robustness against intentional attacks alongside minimal wrong detections which establishes real-time operational trustworthiness for contemporary security systems. The classification performance evaluation matrix presents outcomes for various attack categories simultaneously shown in Figure 10. The developed model achieves great precision and recall levels across all categories while misidentifying a small number of difficult to detect classes primarily Infiltration and Botnet.

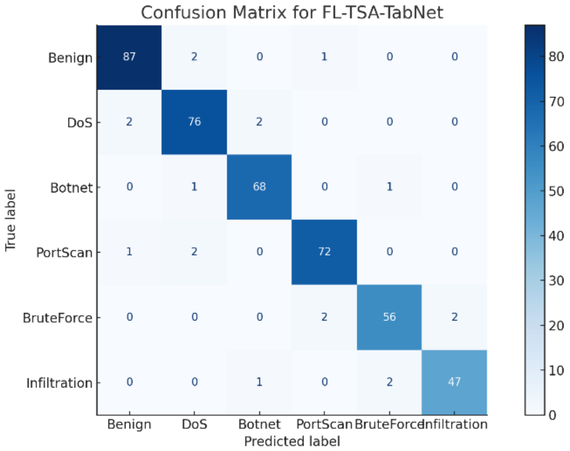

Fig 10. Confusion Matrix for FL-TSA-TabNet

TABLE VI. COMPARATIVE ANALYSIS OF FL-TSA-TABNET WITH EXISTING IDS METHODOLOGIES

Model / Approach	Temporal Modeling	Feature Selection Strategy	Privacy Preservation	Accuracy (%)
CNN-LSTM [12]	LSTM	Manual or None	No	95.02
Dugat-LSTM [18]	LSTM	KPCA + CHA-HBO	No	96.76
M-MultiSVM [16]	No	ASmoT + M-Svd	No	97.89
FL-based IDS [8]	No	None	Yes	95.64
CNN-GRU [2]	GRU	Basic Filtering	No	96.38
ResNet-BiLSTM [3]	BiLSTM	None	No	96.71
FL-TSA-TabNet (Proposed)	TSA Module	Dual-Stage Hybrid Selector (DSHS)	Federated Learning	**97.83**

Table 6 and Figure 11 shows a comparative report of FL-TSA-TabNet with the existing intrusion detection methodologies with respect to their temporal modeling, feature selection, privacy preservation, and classification accuracy. Such traditional models as CNN-LSTM and ResNet-BiLSTM utilize temporal features but have neither strong feature selection nor privacy-sensitive training. The cross-breed models, like Dugat-LSTM and M-MultiSVM, uses feature selection methods like KPCA and ASmoT with the advantage of being centralized.

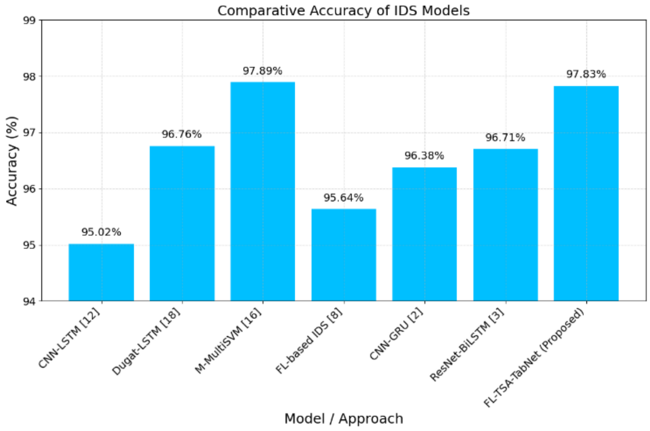

Fig 11. Comparative Accuracy of IDS Models

Although FL-based IDS is privacy preserving, it does not provide any temporal modeling or interpretability. FL-TSA-TabNet integrates Temporal Self-Attention (TSA), a Dual-Stage Hybrid Selector (DSHS), and federated learning to an otherwise unseen extent, thus reaching an impressive 97.83% accuracy mark and surpassing or at least equalling all the baselines in point of crucial criteria.

V. Conclusion And Future Scope

The presented framework FL-TSA-TabNet solves traditional detection system problems through its novel approach which incorporates TabNet architecture together with Temporal Self-Attention. The upgrade of TabNet with Temporal Self-Attention in its core identifies crucial features and discovers temporal connections inside network traffic data. The Dual-Stage Hybrid Selector (DSHS) implements ReliefF and SHAP analysis to create a new feature selection method which reduces redundancy and selects high-importance attributes for improved functionality. A decentralized model training implementation within federated learning protocols protects user data privacy and delivers strong global detection solutions. The proposed model performed excellently against state-of-the-art models in CSE-CIC-IDS2018 dataset evaluations demonstrating a classification accuracy at 97.83% surpassing CNN-GRU (96.38%) and ResNet-BiLSTM (96.71%). The operational readiness of real-time network systems benefits significantly from FL-TSA-TabNet because it achieves quick inference processing together with better adversarial protection mechanisms and additional interpretability features. The model design fits perfectly into edge computing frameworks and cyber-physical systems where data consolidation at a central location becomes impossible. Future modifications to the proposed framework should enable it to handle cross-dataset intrusion data while learning from different intrusion datasets including UNSW-NB15 or TON_IoT through transfer learning techniques or domain adaptation methods. There are promising advancements including blockchain-based secure aggregation techniques and reinforcement learning agents which operate through adaptive threat mitigations. FL-TSA-TabNet represents a vital foundational element which combines explainable capabilities with data protection along with advanced prediction abilities for the development of advanced future-generation cybersecurity technologies.

References

[1] Malothu, A., Kumar, B. S., Prasad, P. R. K., Reddy, S. K., & Reddy, T. S. (2024). Network intrusion detection system by applying ensemble model for smart home. International Journal of Electrical and Computer Engineering, 14(3), 3485–3494. https://doi.org/10.11591/ijece.v14i3.pp3485-3494

[2] Gunupudi, S. C. K., Reddy, G. V. P., Mallela, K., & Allam, A. R. (2024). Deep residual convolutional neural network: An efficient technique for intrusion detection system. Expert Systems with Applications, 238(Part B), 121912. https://doi.org/10.1016/j.eswa.2023.121912

[3] Hamdi, N., Kherchi, I., Ezzedine, H., Benali, Y., & Derdouri, L. (2025). A hybrid learning technique for intrusion detection system for smart grid. Sustainable Computing: Informatics and Systems, 46, 101102. https://doi.org/10.1016/j.suscom.2025.101102

[4] Alotaibi, M., Bhatnagar, S., Srivastava, R., & Jamal, S. S. (2025). Hybrid GWQBBA model for optimized classification of attacks in intrusion detection system. Alexandria Engineering Journal, 116, 9–19. https://doi.org/10.1016/j.aej.2024.12.057

[5] Kumar, V., Sharma, S., Arora, A., & Rai, H. (2025). NIDS-DA: Detecting functionally preserved adversarial examples for network intrusion detection system using deep autoencoders. Expert Systems with Applications, 270, 126513. https://doi.org/10.1016/j.eswa.2025.126513

[6] Serrano, W., Dutt, M., Al-Mousa, A., Karki, M., & O'Connell, T. (2025). CyberAIBot: Artificial intelligence in an intrusion detection system for cybersecurity in the IoT. Future Generation Computer Systems, 166, 107543. https://doi.org/10.1016/j.future.2024.107543

[7] Rajathi, C., Krishnan, R., Suganthi, M., & Arulmurugan, R. (2025). Hybrid learning model for intrusion detection system: A combination of parametric and non-parametric classifiers. Alexandria Engineering Journal, 112, 384–396. https://doi.org/10.1016/j.aej.2024.10.101

[8] Zukaib, U., Asghar, M. N., Ahmed, M., Javed, A. R., & Abid, A. (2025). Mitigating backdoor attacks in federated learning based intrusion detection systems through neuron synaptic weight adjustment. Knowledge-Based Systems, 314, 113167. https://doi.org/10.1016/j.knosys.2025.113167

[9] Wu, X., Zhang, L., Yang, B., Chen, Y., & Liu, S. (2025). Boosting incremental intrusion detection system with adversarial samples. Expert Systems with Applications, 271, 126632. https://doi.org/10.1016/j.eswa.2025.126632

[10] Berguiga, A., Laassiri, J., Rahmani, M., & Hmina, N. (2025). HIDS-IoMT: A deep learning-based intelligent intrusion detection system for the Internet of Medical Things. IEEE Access, 13, 32863–32882. https://doi.org/10.1109/ACCESS.2025.3543127

[11] Mahmoud, M. M., Badran, M. F., Darwish, A., & Elsisi, M. (2025). XI2S-IDS: An explainable intelligent 2-stage intrusion detection system. Future Internet, 17(1), 25. https://doi.org/10.3390/fi17010025

[12] Bamber, S. S., Kaur, H., Arora, A., & Singh, M. (2025). A hybrid CNN-LSTM approach for intelligent cyber intrusion detection system. Computers & Security, 148, 104146. https://doi.org/10.1016/j.cose.2024.104146

[13] Kaushik, S., Singh, A., Gupta, D., & Jain, R. (2025). Robust machine learning based intrusion detection system using simple statistical techniques in feature selection. Scientific Reports, 15, 3970. https://doi.org/10.1038/s41598-025-88286-9

[14] Tahir, M., Irfan, M., Saleem, K., & Hussain, A. (2025). A novel approach for handling missing data to enhance network intrusion detection system. Journal of Computational Science Advances, 3, 100063. https://doi.org/10.1016/j.csa.2024.100063

[15] Saravanan, S., Thangavel, K., Elayaraja, M., & Rajasekaran, M. P. (2025). Deep learning models for intrusion detection systems in MANETs: A comparative analysis. Data Management and Analytics, 3(1), 96–110. https://doi.org/10.31181/dma31202556

[16] Turukmane, A. V., Jadhav, B., Dhabu, M., & Rane, P. (2024). M-MultiSVM: An efficient feature selection assisted network intrusion detection system using machine learning. Computers & Security, 137, 103587. https://doi.org/10.1016/j.cose.2023.103587

[17] Nandanwar, H., Sharma, N., Kapoor, R., & Dwivedi, V. (2024). Deep learning enabled intrusion detection system for industrial IoT environment. Expert Systems with Applications, 249(Part C), 123808. https://doi.org/10.1016/j.eswa.2024.123808

[18] Devendiran, R., Ilango, V., Rajesh, M., & Sivakumar, R. (2024). Dugat-LSTM: Deep learning based network intrusion detection system using chaotic optimization strategy. Expert Systems with Applications, 245, 123027. https://doi.org/10.1016/j.eswa.2023.123027

[19] Li, J., Sun, H., Zhang, W., Wang, Q., & Huang, F. (2024). Optimizing IoT intrusion detection system: Feature selection versus feature extraction in machine learning. Journal of Big Data, 11, 36. https://doi.org/10.1186/s40537-024-00892-y

[20] Sadia, H., Shabbir, M., Qayyum, A., Khan, Z. H., & Riaz, M. (2024). Intrusion detection system for wireless sensor networks: A machine learning based approach. IEEE Access, 12, 52565–52582. https://doi.org/10.1109/ACCESS.2024.3380014

[21] https://www.unb.ca/cic/datasets/ids-2018.html accessed on 17th March 2025.