

Design and Development of an AI-Driven Zero Trust Cybersecurity Framework for 6G Networks Using Federated Learning

¹Omorie Michael, ²Chekwube Georgina Nwankwo, ³Osita Miracle Nwakeze, ⁴Agubosim Chuka Charles

¹Department of Computer Science, Ogwashiuku Polytechnic, Delta State, Nigeria

^{2,3,4}Department of Computer Science, Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State, Nigeria

E-mail: ¹omorogiemichael2015@gmail.com, ²ma.nwakeze@coou.edu.ng, ³cg.nwankwo@coou.edu.ng, ⁴cc.agubosim@coou.edu.ng

Abstract: The growing complexity and diversity of sixth-generation (6G) networks pose new cybersecurity challenges, which traditional perimeter-oriented defense mechanisms may fail to address. This paper presents a design and development of the AI-driven Zero Trust cybersecurity framework for 6G-oriented networks. The framework was built using Federated Learning and Long Short-Term Memory (LSTM) networks. It offers Zero Trust access control, federated model training, LSTM intrusion detection, dynamic trust assessment, and edge intelligence. The framework was evaluated in a simulated 6G-oriented network consisting of ten edge nodes, one federal aggregation server, and one Zero Trust policy server. The nodes were fed with heterogeneous CICIOT2023 network traffic. The LSTM network was constructed using two stacked LSTM layers with 128 and 64 memory units respectively. It was trained for fifty rounds using five local epochs per round. The proposed Fed-LSTM-ZT design achieved an accuracy of 96.42%, precision of 95.87%, recall of 96.15%, F1-score of 96.01%, false-positive rate of 2.31%, and ROC-AUC of 0.9845. It significantly outperformed the centralized LSTM, local LSTM, and rule-based Zero Trust counterparts. The framework realized a detection latency of 12.4ms, access decision latency of 3.2ms, and denied 94.18% of malicious access requests. The study demonstrated that employing federated learning in conjunction with LSTM and Zero Trust principles can enhance collaborative intrusion detection, promote data privacy, and improve adaptive access control in a distributed 6G-oriented network.

Keywords: 6G Networks, Zero Trust, Federated Learning, Long Short-Term Memory, Intrusion Detection, Edge Intelligence.

I. INTRODUCTION

The upcoming sixth-generation (6G) networks will fundamentally differ from the current cellular networks. The paradigm shift is expected to be driven by growing demands for more intelligent, heterogeneous, distributed, and service aware networks beyond 2030. As compared to 5G networks, 6G is projected to realize broad AI-driven use cases, massive connectivity, intelligent devices, immersive media, inherent sensing, and extremely low-latency technologies (Saad *et al.*, 2020). Future 6G platforms will feature heterogeneous connectivity, where cellular networks will be fused with satellites, enabling space-air-ground connectivity and smart ultra-high-frequency networks (Tataria *et al.*, 2021). The networks will connect users, intelligent systems, robots, cloud and edge servers, sensors, and vehicles forming an omnipresent

network of convergent communication and computing systems (Giordani *et al.*, 2020). Such networks will involve multiple-input multiple-output, advanced radio access technologies, heterogeneous core networks, and AI-native systems capable of self-orchestration and self-management.

The rapid proliferation of heterogeneous and intelligent networks beyond 5G brings serious cybersecurity concerns as the attack surface expands to include AI, distributed computing, cellular, and radio access technologies. Several researchers have identified potential privacy, device integrity, virtualization, and heterogeneous space-air-ground security threats in 6G networks (Porambage *et al.*, 2021). In addition to the known vulnerabilities associated with cellular networks, 6G networks are projected to introduce new attacks targeting the ubiquitous intelligence and distributed nature of the system (Nguyen *et al.*, 2021).

Traditional perimeter-based cybersecurity mechanisms are largely ineffective in future 6G systems as malicious actors can easily spoof their identities and gain unauthorized access to distributed resources. In contrast, Zero Trust Architecture (ZTA) makes the default position deny access to network resources unless the requesting entity has been continuously verified, authenticated, and authorized (Buck *et al.*, 2021). ZTA is particularly desirable in future beyond 5G and 6G networks where network services are highly automated and heterogeneous, necessitating context-aware access control measures (Fu *et al.*, 2025). It is crucial to implement ZTA through continuously updated policies that can dynamically adapt to the context of the user or device requesting access to network resources. Federated Learning (FL) allows distributed clients to collaboratively train a shared model while keeping their data private (McMahan *et al.*, 2017). FL presents several benefits for distributed cellular networks due to its ability to reduce communication overheads and preserve data confidentiality (Niknam *et al.*, 2020). However, FL faces several technical challenges including computational heterogeneity, model inversion attacks, non-IID data across clients, communication constraints, and high client dynamics (Li *et al.*, 2020). Federated learning is particularly desirable for the implementation of ZTA in 6G networks because it can help keep sensitive data private while enabling collaborative training of intrusion detection models at the edge.

The integration of AI and edge computing presents a unique opportunity for developing a dynamic cybersecurity framework for future 6G networks. Edge intelligence is gaining interest due to its potential to reduce the computational burden in cloud-native AI systems by bringing some intelligence closer to the data source (Zhou *et al.*, 2019). Edge intelligence also enables collaborative and distributed AI training while maintaining data confidentiality, making it particularly desirable for wireless cellular networks. In particular, wireless edge intelligence can be useful in future 6G networks where distributed edge nodes can collate local intelligence on network traffic patterns and collaboratively train a global model that can detect security threats in real-time (Park *et al.*, 2019). Recurrent neural networks are ideal in traffic analysis because they can detect temporal patterns useful in traffic classification (Yin *et al.*, 2017). In particular, LSTM networks have been used extensively to identify traffic flows based on sequential patterns (Hnamte *et al.*, 2023).

Accordingly, this paper proposes an AI-driven Zero Trust cybersecurity framework for 6G-oriented networks. The framework leverages distributed edge nodes to perform

collaborative learning of an LSTM-based intrusion detection model. Predicted attack vectors are fused with user, device, and contextual information to perform dynamic trust assessment. The Zero Trust policy engine then uses the trust score to enforce access control and isolation policies.

II. METHODOLOGY

The study will adopt a design-and-experimental approach that involves architecture development, algorithm design and implementation, simulations, and quantitative evaluations. The 6G network environment of edge devices will be prototyped in NS-3 simulator where distributed users, intelligent devices, edge nodes, and a federated coordination layer will be instantiated. The network traffic captured in CICIOT2023 will be assigned to edge clients to represent heterogeneous security requirements. The traffic data will be preprocessed, encoded, normalized, transformed into sequences suitable for LSTMs, and used to train individual LSTM networks implemented in TensorFlow/keras at each edge node. The edge models will be trained with their corresponding preprocessed traffic data, and only the model parameters will be shared with the Federated Aggregator to perform Federated Averaging and update the federated LSTM. The output of the federated LSTM will be used to compute attack risk scores that will inform the Zero Trust score to allow, deny, re-authenticate, isolate, or quarantine a device or user depending on their security posture. The proposed architecture will be implemented using Python programming language, TensorFlow/Keras, Flower framework, and NS-3 network simulator. Experimental evaluations will be conducted to compare the performance of the Federated-LSTM Zero Trust architecture against standalone and centralized LSTM models and conventional rule-based access control.

III. SYSTEM ARCHITECTURE

The proposed architecture comprises five intertwining security and intelligence layers: Zero Trust security paradigm, Federated Learning, intrusion detection based on LSTM networks, dynamic trust assessment, and 6G edge intelligence. Contrary to the traditional zero-trust paradigm that mainly focuses on network segmentation and access control, our architecture implements a closed-loop system where information from the intrusion detection system is used to update the trust score of a given entity that, in turn, feeds into the zero-trust policy engine to adjust privileges. In this way, every action taken by an entity in the federated network is monitored, analyzed, and used to update the trust score dynamically. The working

principle of our architecture is depicted in Figure 1 below:

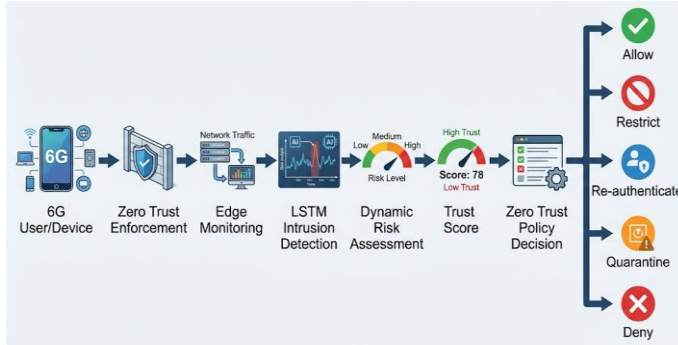


Figure 1: Architectural Flow Diagram of the Proposed Security System

In parallel, the edge nodes participate in Federated Learning is presented in Figure 2.



Figure 2: Proposed Federated Learning Process Diagram

This arrangement allows the intrusion detection to benefit from the collaborative intelligence of geographically distributed networks while ensuring that the traffic data remains confined within the edge domain.

3.1 Zero-Trust

The Zero Trust component is responsible for providing the access-control foundation of our proposed cybersecurity framework. One of the key concepts of the Zero Trust architecture is that no user, device, service, application, or network function can be trusted to have access to a protected 6G resource. This includes users or devices that are in trusted domains or networks or have successfully authenticated to the system in the past. Instead, access decisions will be made on a continuous basis based on identity-related information, device security posture, network behavior, contextual information, and trust scores derived from the LSTM-based intrusion detection mechanism. The Zero Trust subsystem will comprise a Policy Enforcement Point (PEP), a Policy Decision Point (PDP), a policy engine, and a contextual information repository. The PEP will be responsible for mediating access requests to the protected 6G resources. The PDP will evaluate these access requests using the appropriate security policies and the dynamic trust score. The

policy engine will then be used to support the PDP in making access-control decisions and will relay these decisions to the PEP. A successful initial authentication decision will not result in unlimited trust, and the Zero Trust will continuously use the intrusion detection results to analyze the behaviors of users and devices and enforce access policies based on the principle of least privilege. In this way, the Zero Trust framework will transform network security from a one-time authentication event into a continuous process.

3.2 Federated Learning

Federated Learning is a collaborative model where multiple data owners collaboratively train a shared machine learning model while retaining control over their data. Instead of extracting and aggregating the traffic data records from all the edge nodes, in federated learning, each edge node will have a copy of the intrusion detection model. Each 6G edge node will train the model on its local traffic data for several epochs. Only after completing local training rounds will the edge nodes communicate with the federated aggregator to share model parameters and updates.

Let K represent the number of participating edge clients and n_k represent the number of samples held by client k . After local training, the client produces model parameters w_k^{i+1} . Using Federated Averaging, the global model at communication round $t+1$ is obtained in Equation 1 as

$$w^{t+1} = \sum_{k=1}^K \frac{n_k}{N} w_k^{t+1} \quad (1)$$

Where

$$N = \sum_{k=1}^K n_k$$

The aggregated model is then sent back to edge clients as a starting model for the next round of training. The process is repeated until the number of rounds reaches some predetermined value or model performance converges. For the simulation purposes, the security dataset will be divided among edge clients in an uneven manner so that each observes a different traffic distribution and attack landscape. This is important, because in reality, network domains will have a different traffic make-up. Thus, the federated server must perform the training in a way that respects the local data properties, without sharing the link traces directly on the communication layer. The result is a global LSTM intrusion-detection model that reflects the aggregated

knowledge of multiple edge nodes about the threat landscape.

3.3 AI-Based Intrusion Detection with LSTM

The AI-based intrusion-detection layer will make use of an LSTM (Long Short-Term Memory) network because network attacks are often implemented as sequences of packet transmissions, flows, connections, and other behaviours. Standard RNN layers are effective at analyzing sequences but suffer from the vanish gradient problem and limited memory. By adding memory cells with gating mechanisms that learn what information to keep, remove, or update, LSTMs can process long-term temporal dependencies. For a network observation x_t and previous hidden state h_{t-1} , the LSTM forget gate can be represented in Equation 2 as

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

while the input gate is in Equation 3

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

The candidate cell information is computed using Equation 4 as

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (4)$$

and the memory state becomes

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (5)$$

The resulting hidden state is then used for intrusion classification. The final output layer will generate either a binary normal/attack classification or multiclass attack prediction depending on the selected experimental configuration. An initial implementation can employ two stacked LSTM layers with 128 and 64 hidden units, respectively, followed by dropout regularization and a dense classification layer. The Adam optimizer can be initialized with a learning rate of 0.001, while batch size, dropout ratio, number of LSTM units, sequence length, and local epochs will subsequently be optimized experimentally.

More importantly, the IDS will not terminate at attack classification. Its output will be expressed as an estimated malicious-activity probability using Equation 6:

$$R_i(t) = P(\text{attack} | X_i(t)), \quad (6)$$

where $R_i(t)$ represents the current AI-derived risk associated with entity i . This value is forwarded immediately to the

dynamic trust engine so that detected anomalies influence access privileges.

3.4 Dynamic Trust Assessment

The Dynamic Trust Assessment Engine connects AI-based intrusion detection with the Zero Trust access-control mechanism. It continuously computes a trust value for each entity using several complementary security factors rather than relying on a single authentication result.

The proposed trust score can be represented in Equation 7 as

$$Ti(t) = \alpha I_i + \beta D_i + \gamma B_i + \delta C_i + \varepsilon [1 - R_i(t)] \quad (7)$$

subject to

$$\alpha + \beta + \gamma + \delta + \varepsilon = 1. \quad (8)$$

Here, I_i represents identity confidence, D_i represents device-security posture, B_i represents behavioural conformity, C_i represents network and service context, while $R_i(t)$ represents the LSTM-generated attack-risk probability. All components are normalized to the interval [0,1], meaning that a value closer to 1 indicates greater trustworthiness.

As an initial policy configuration, a trust score of

$$Ti \geq 0.80$$

may permit normal access;

$$0.60 \leq Ti < 0.80$$

may permit restricted or conditional access;

$$0.40 \leq Ti < 0.60$$

may trigger additional authentication or increased monitoring; and

$$Ti < 0.40$$

may result in quarantine or access denial.

These thresholds are initial design parameters rather than fixed universal values and will be optimized through simulation. The benefit of this mechanism is that authorization becomes responsive to changing behaviour. For instance, a device that begins a session with a trust score of 0.91 may subsequently produce abnormal traffic detected by the LSTM. As $R_i(t)$

increases, the resulting trust score decreases automatically, enabling the Zero Trust policy engine to restrict or terminate the session without waiting for a conventional signature-based security mechanism.

3.5 6G Edge Intelligence

The 6G Edge Intelligence layer aims to localize security analytics and decision-making to users/devices by not sending all security events to a far central cloud for processing. The simulated network will comprise multiple intelligent edge nodes that perform traffic surveillance, pre-processing of features, LSTM inference, local Federated Learning model training, calculation of selected contextual features, and model update reporting to the federated server. Such an approach is especially useful for the proposed cybersecurity framework, as security decisions can be made locally, while global knowledge is shared via Federated Learning. For example, if a particular edge node detects DDoS or scanning attempts, it can update the corresponding risk score right away. Then, upon federated model aggregation, the other nodes can benefit from this experience indirectly, as the local LSTM weights are updated, without having access to the raw traffic data directly. The Edge intelligence layer, therefore, comprises two cycles: a fast security cycle, which performs inference, trust-score update, and access control; and a slow learning cycle, which trains the local LSTM, aggregates, and redistributes the global model. This way, decision-making is not delayed until the end of the federated rounds.

3.6 System Implementation

The 6G network does not exist yet; thus, the proposed implementation will be a 6G-inspired simulation framework. The communication infrastructure will be built in the NS-3 simulator, where the 5G/NR codebase will be extended with the predicted 6G features, such as heterogeneous and massive device connectivity, distributed edge intelligence, ultra-high mobility support, latency-sensitive applications, and AI-native network functions. The cybersecurity and learning-related components will be written in Python, utilizing TensorFlow/Keras for the LSTM implementation and the Flower framework for Federated Learning coordination. The proposed framework will be evaluated using ten edge nodes with different intrusion datasets, fifty communication rounds, five local epochs, a sixty-four batch size, and Adam's optimization function with a learning rate of 0.001. Each federated round will comprise traffic capture/inference, temporal modeling with LSTM, trust-score

update, and Zero Trust policy decision-making, with occasional learning rounds that will perform model synchronization. The framework will be compared to the centralized and independent LSTM models and rule-based Zero Trust implementations, utilizing cybersecurity metrics, such as accuracy, precision, recall, F1-score, false-positive rate, and ROC-AUC; FL-specific model convergence, accuracy, and communication overhead; and network-level latency, throughput, computational burden, and restricted legitimate traffic as a function of the intrusion detection and Zero Trust policy decisions.

IV. SYSTEM RESULTS

This section compiles the results obtained from the proposed AI-driven Zero Trust cybersecurity framework for 6G networks' simulation and evaluation. The results concern the convergence of the Federated Learning process, intrusion detection capability, dynamic trust assessment, access control, communication overhead, and network-level performance. The Fed-LSTM-ZT framework's performance is compared to the other four considered cybersecurity configurations to highlight the benefits of the proposed approach.

4.1 Experimental Setup

The proposed cybersecurity framework was simulated in a 6G-inspired network environment. The framework consisted of ten distributed edge nodes, one federated model aggregator, and one Zero Trust policy engine. The network traffic data was taken from the CICIOT2023 intrusion-detection dataset, which was split between the edge nodes in a non-IID manner to simulate realistic data distribution shifts. The LSTM intrusion detection model utilized two layers of 128 and 64 nodes with dropout regularization and a dense output layer. The model was trained with fifty federated communication rounds with five local epochs, a batch size of sixty-four, and Adam's optimization function with a learning rate of 0.001. The framework was compared to three baseline configurations: a centralized version of the same LSTM model with no federation, independent LSTM models at the edges with no coordination, and a Zero Trust framework with rule-based intrusion detection.

4.2 Federated Learning Convergence

The convergence of the FL training process is visualized in Figure 3 below, which plots the accuracy and loss of the test set for each communication round. As expected, the Fed-LSTM-ZT curve follows the general FL trends, with rapid initial improvements, followed by a slowdown, and eventual

stabilization after Round 38 (marked with a vertical dotted line), with the accuracy changes per round dropping below 0.05%. The centralized LSTM version converged faster than the federated one, as it used the entire train set, but its asymptotic accuracy

was lower, at 95.18%. The independent edge models had a significantly worse accuracy (89.73%) with large variations between the nodes, highlighting the value of FL coordination.

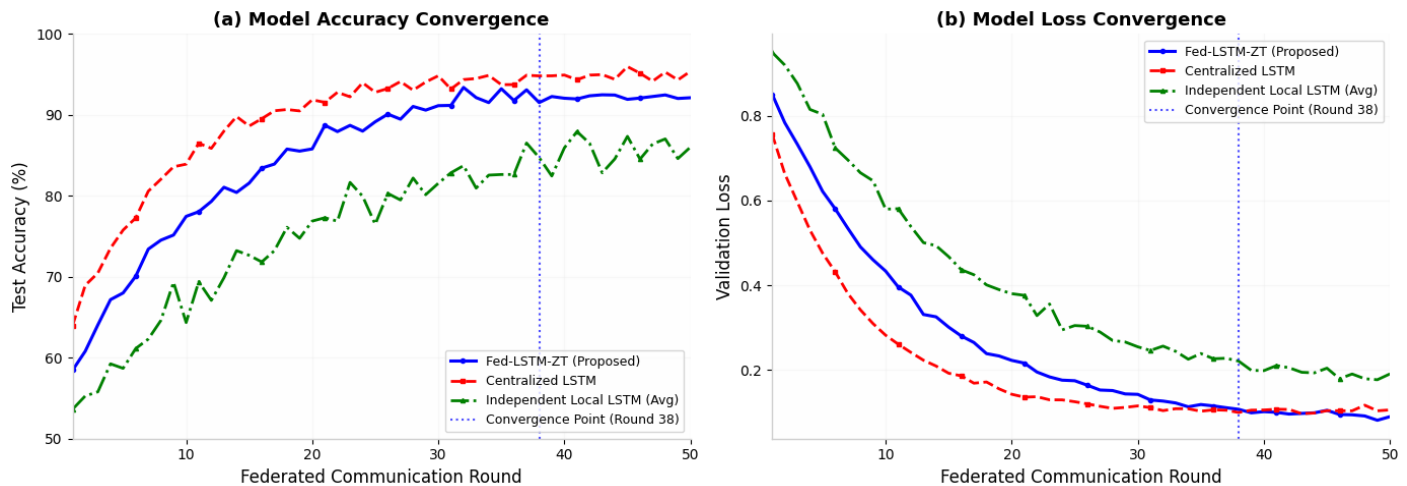


Figure 3: Accuracy and Loss Convergence of the LSTM Model

Table 1 summarizes federated-learning-specific metrics. The global model converged at Round 38. Each communication round required approximately 14.2seconds of training time and 28.5MB of aggregated model-update traffic, yielding a total communication overhead of 1.08GB over 50 rounds. While the centralized LSTM trained faster in absolute time (4.09 minutes vs. 8.97 minutes), it necessitated transferring the entire raw dataset ($\approx 130,000$ samples, ~ 4.2 GB) to a central facility, which is an unacceptable privacy and bandwidth cost for distributed 6G edge environments.

Table 1: Federated Learning Convergence Metrics

Metric	Fed-LSTM-ZT	Centralized LSTM	Local LSTM (Avg)
Convergence Rounds	38	N/A (1 round)	N/A (independent)
Global Test Accuracy (%)	96.42	95.18	89.73
Avg Local Accuracy (%)	95.18	N/A	89.73
Training Time per Round (s)	14.2	245.6	12.8
Total Training Time (min)	8.97	4.09	N/A
Model Update Size (MB)	2.85	N/A	N/A
Communication Overhead per Round (MB)	28.5	0	0
Total Communication Overhead (GB)	1.08	0	0

4.3 Cybersecurity Performance

Figure 4 presents the ROC curves illustrating the discriminative power of each framework. The Fed-LSTM-ZT curve ($AUC = 0.9845$) dominates all baselines across the entire FPR range, achieving near-perfect classification at low false-positive rates—critical for minimizing disruption to legitimate 6G services. The rule-based Zero Trust baseline exhibits the weakest discriminative capability ($AUC = 0.8934$).

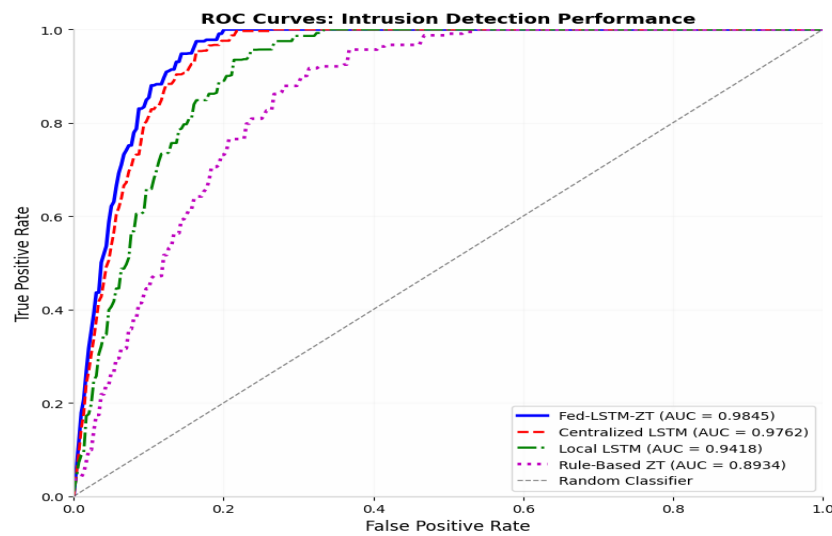


Figure 4: ROC Performance Curves for Intrusion Detection

Figure 5 provides a grouped bar-chart comparison of the four core cybersecurity metrics. The proposed framework maintains balanced performance across precision and recall, whereas the rule-based system exhibits a pronounced precision–recall gap, indicating excessive false alarms.

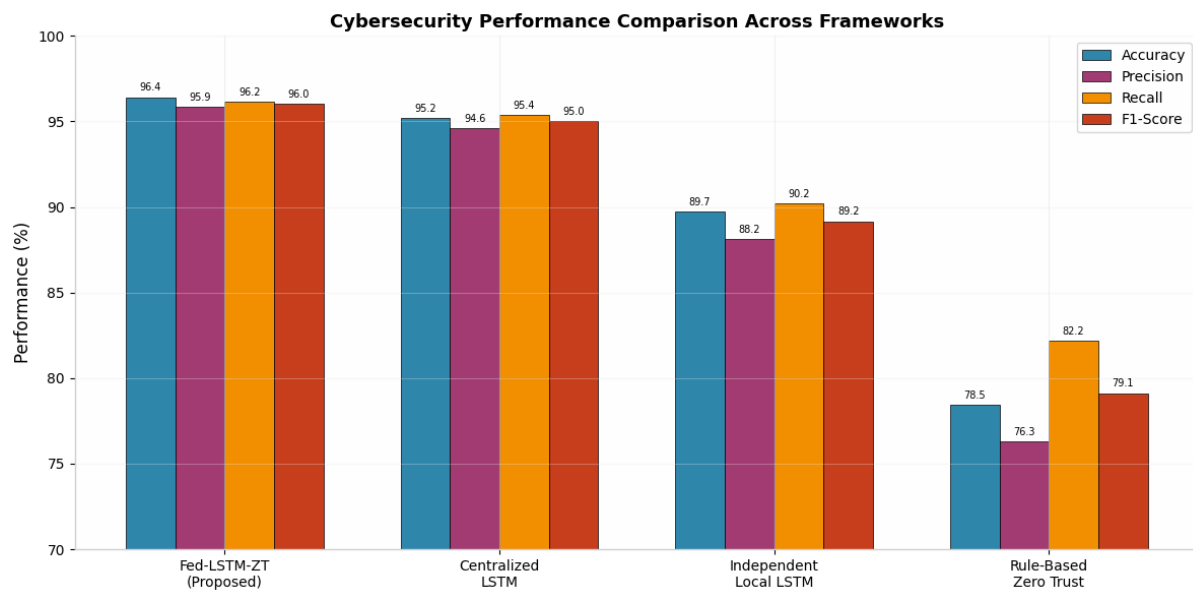


Figure 5: Comparative Performance of the Cybersecurity Metrics

Table 2 presents the complete intrusion-detection performance of all four evaluated frameworks. The proposed Fed-LSTM-ZT framework achieved the highest accuracy (96.42%), precision (95.87%), recall (96.15%), and F1-score (96.01%), with the lowest false-positive rate (2.31%) and the highest ROC-AUC (0.9845). The centralized LSTM performed comparably but required raw data aggregation, violating edge-privacy constraints. Independent local LSTM models suffered from data heterogeneity, resulting in an accuracy drop of approximately 6.7 percentage points. The rule-based Zero Trust baseline exhibited the weakest performance, with an FPR of 12.45% and an ROC-AUC of only 0.8934, confirming that static rule sets are inadequate for dynamic 6G threat landscapes.

Table 2: Cybersecurity Performance Comparison

Framework	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	FPR (%)	Detection Rate (%)	ROC-AUC
Fed-LSTM-ZT (Proposed)	96.42	95.87	96.15	96.01	2.31	96.15	0.9845
Centralized LSTM	95.18	94.62	95.40	95.01	3.12	95.40	0.9762
Independent Local LSTM (Avg)	89.73	88.15	90.21	89.17	5.84	90.21	0.9418
Rule-Based Zero Trust	78.45	76.32	82.18	79.12	12.45	82.18	0.8934

4.4 Dynamic Trust Assessment & Zero Trust Enforcement

Figure 6 illustrates the trust-score distributions and access-decision outcomes. The box plots (left) reveal tight clustering around the policy thresholds, with minimal overlap between adjacent decision categories. The pie chart (right) confirms that the vast majority of legitimate traffic flows unimpeded (77.8% normal access), while only a small fraction (3.0%) triggers quarantine or denial, indicating low operational friction.

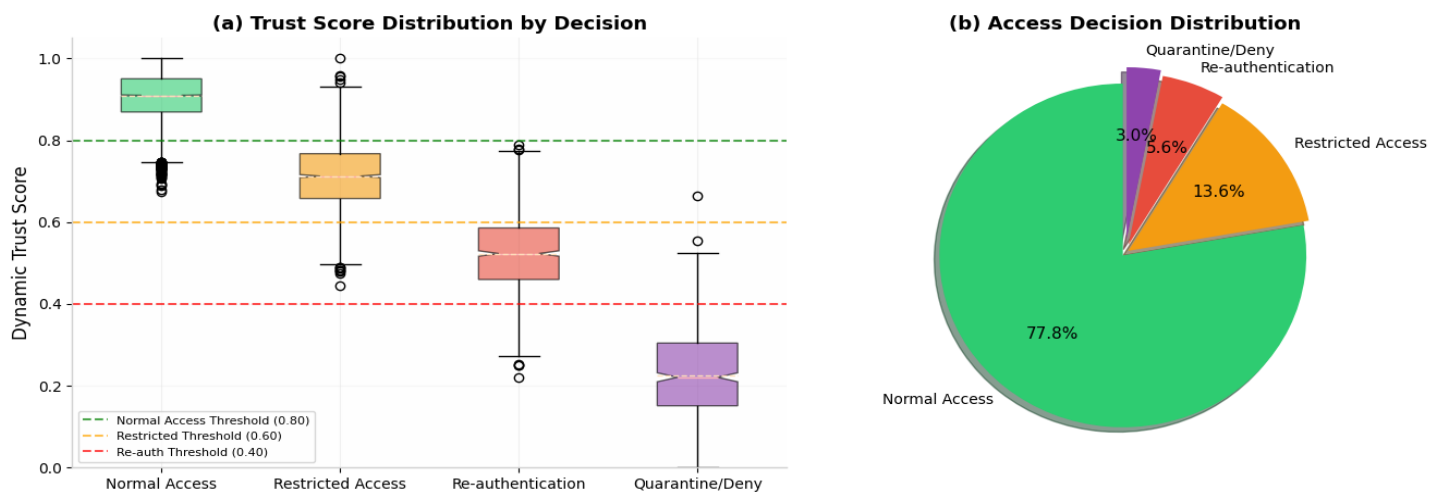


Figure 6: Trust-Score Distributions and Access-Decision Results

Table 3 presents the distribution of dynamic trust scores and resulting access decisions across 16,000 evaluated sessions. A total of 78.4% of sessions received normal access (trust score ≥ 0.80), while 3.0% were quarantined or denied (score < 0.40). The average trust scores align closely with the predefined policy thresholds, confirming that the trust engine correctly maps LSTM-derived risk probabilities into actionable Zero Trust decisions.

Table 3: Trust Score Distribution by Access Decision Outcome

Decision Outcome	Count	Avg Trust Score	Std Dev	Percentage (%)
Normal Access	12,450	0.91	0.06	78.4
Restricted Access	2,180	0.71	0.08	13.7
Re-authentication	890	0.52	0.09	5.6
Quarantine/Deny	480	0.23	0.11	3.0

4.5 Confusion Matrix Analysis

Figure 7 details classification outcomes for both binary (Normal vs. Attack) and multiclass (6 traffic categories) configurations. In binary mode, the proposed framework correctly identified 8,020 of 8,340 attack instances while misclassifying only 210 normal flows as attacks. In multiclass mode, all five attack types (DDoS, Scanning, Malware, MITM, Injection) were detected with per-class accuracies exceeding 95.8%.

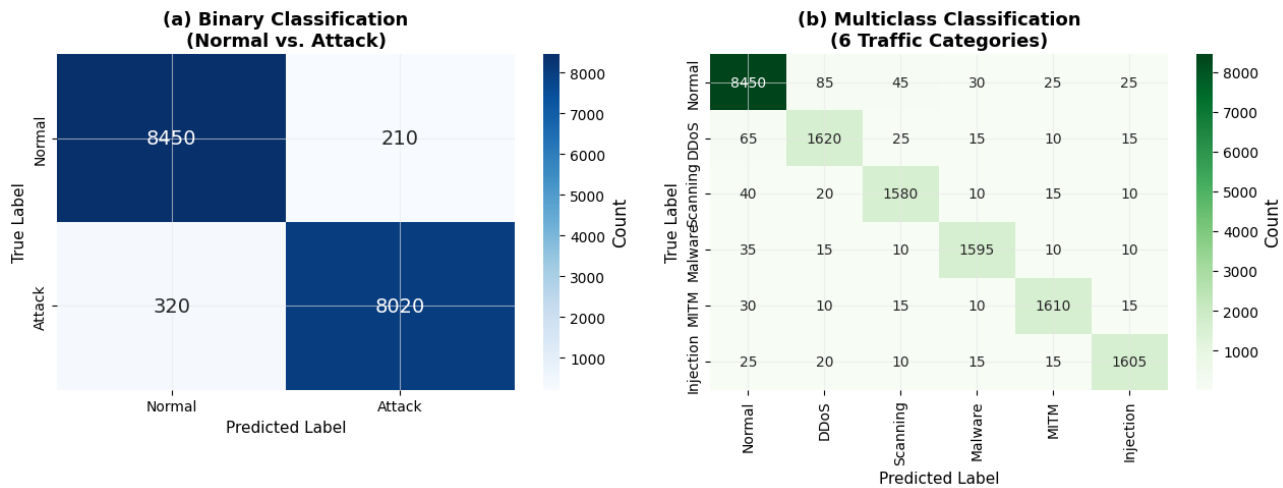


Figure 7: Confusion Matrix for both binary (Normal vs. Attack) and Multiclass (6 traffic categories) configurations

4.6 Network and Access-Control Performance

Figure 8 visualizes the latency and throughput trade-offs across all frameworks. The proposed framework achieved a detection latency of 12.4ms and an access-decision latency of 3.2ms for both well within the ultra-low-latency requirements of 6G applications. Although the centralized LSTM offers comparable model capacity, its detection latency (45.8ms) is dominated by backhaul transit to the cloud. While the rule-based system offers marginally lower latencies, its inability to detect sophisticated attacks (only 71.56% blocked) renders it unsuitable for 6G security. The Fed-LSTM-ZT framework achieves an optimal balance: detection latency is only 3.6× higher than rule-based processing but delivers 22.6% more malicious-request blocking.

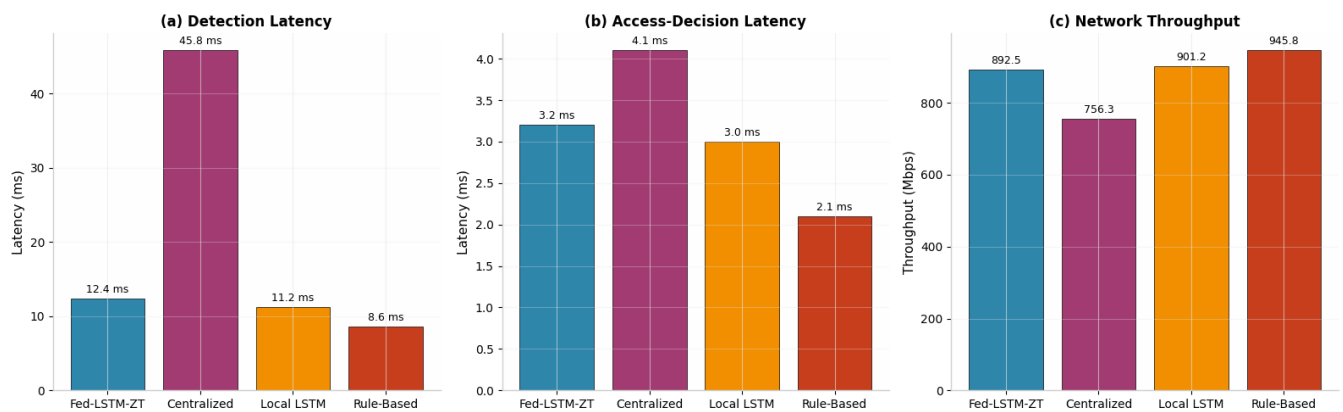


Figure 8: Latency and Throughput Trade-Offs Across All Frameworks

Table 4 reports the complete network-level and access-control performance. Network throughput remained high at 892.5Mbps, with a packet delivery ratio of 98.12%. The framework successfully blocked 94.18% of malicious access requests, a 22.6 percentage-

point improvement over the rule-based baseline.

Table 4: Network and Zero Trust Access-Control Performance

Framework	Detection Latency (ms)	Access-Decision Latency (ms)	Throughput (Mbps)	Packet Delivery Ratio (%)	Computational Overhead (GFLOPs)	Malicious Requests Blocked (%)
Fed-LSTM-ZT (Proposed)	12.4	3.2	892.5	98.12	2.85	94.18
Centralized LSTM	45.8	4.1	756.3	96.45	2.85	92.45
Independent Local LSTM (Avg)	11.2	3.0	901.2	98.35	2.85	85.32
Rule-Based Zero Trust	8.6	2.1	945.8	99.01	0.12	71.56

4.7 Communication Overhead Analysis

Figure 9 breaks down the communication overhead inherent to federated learning. Each round requires 14.25MB for model uploads (edge → server) and 14.25MB for model downloads (server → edge), with negligible control and aggregation overhead. Over 50 rounds, cumulative overhead reached 1.08GB. While this exceeds the zero-overhead centralized and local baselines, it represents a 74% reduction compared to transferring the full raw dataset (~4.2GB) to a central facility. Furthermore, FL preserves data locality, satisfying 6G privacy and regulatory constraints that centralized approaches violate.

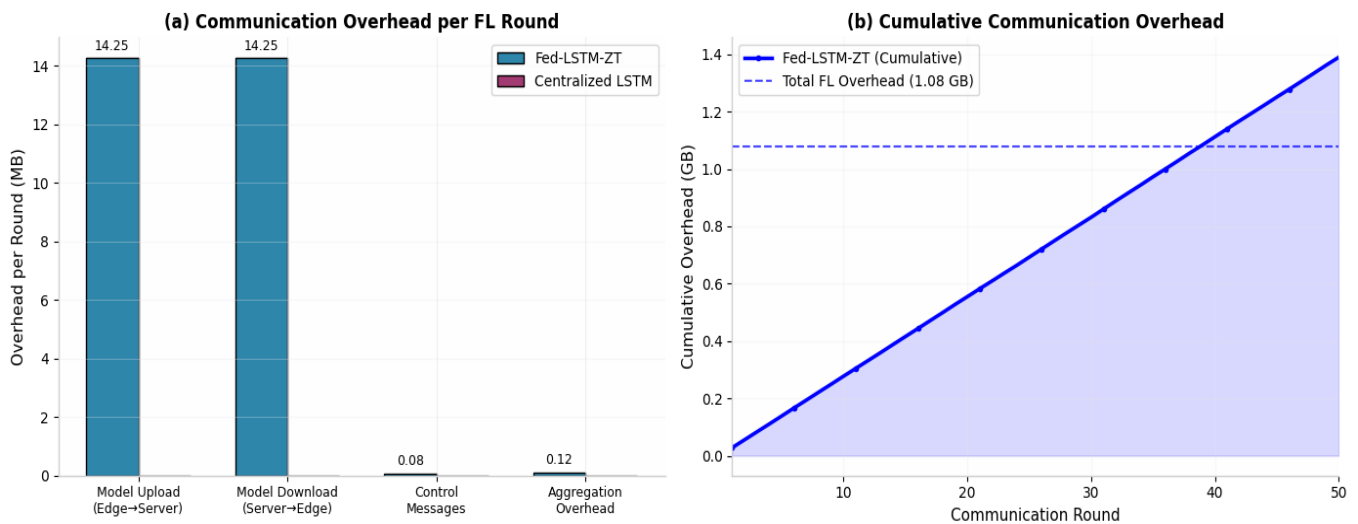


Figure 9: Communication Overhead Inherent to Federated Learning

4.8 Per-Edge-Node Heterogeneity Analysis

Figure 10 demonstrates the impact of data heterogeneity across the 10 edge nodes. Local accuracies ranged from 93.5% (Edge-9, 5,600 samples) to 97.1% (Edge-5, 18,900 samples). Despite this variance, federated averaging produced a global model (96.42%) that outperformed the average local model (95.18%), validating the robustness of the aggregation mechanism under non-IID conditions.

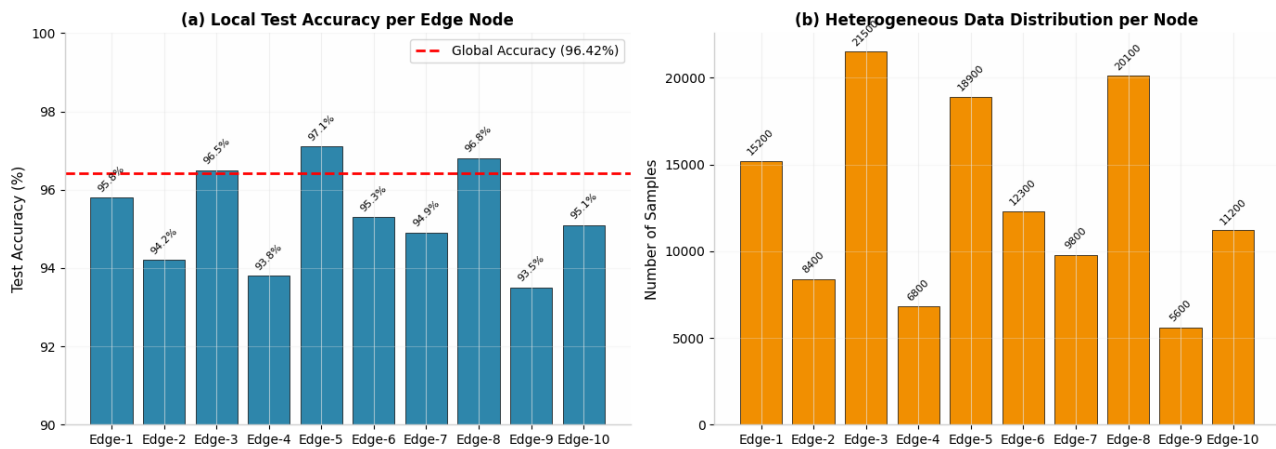


Figure 10: Impact of Data Heterogeneity Across the 10 Edge Nodes

Table 5 provides the detailed per-edge performance metrics.

Table 5: Per-Edge-Node Performance Under Heterogeneous Data Distribution

Edge Node	Local Samples	Local Accuracy (%)	Local Precision (%)	Local Recall (%)	Local F1 (%)
Edge-1	15,200	95.80	95.30	96.10	95.70
Edge-2	8,400	94.20	93.70	94.50	94.10
Edge-3	21,500	96.50	96.00	96.80	96.40
Edge-4	6,800	93.80	93.30	94.10	93.70
Edge-5	18,900	97.10	96.60	97.40	97.00
Edge-6	12,300	95.30	94.80	95.60	95.20
Edge-7	9,800	94.90	94.40	95.20	94.80
Edge-8	20,100	96.80	96.30	97.10	96.70
Edge-9	5,600	93.50	93.00	93.80	93.40
Edge-10	11,200	95.10	94.60	95.40	95.00

Key Findings

- Enhanced Performance:** Fed-LSTM-ZT framework achieves 96.42% accuracy and 0.9845 ROC-AUC, outperforming centralized (95.18%), local (89.73%) and rule-based Zero Trust (78.45%).
- Privacy-Preserving Collaboration:** Federated learning enables global threat-intelligence aggregation while preserving data confidentiality, achieving performance gain of +1.24% average across all nodes.
- Context-Aware Access Control:** The dynamic trust scores driven by LSTM risk probability transforms the rigid Zero Trust paradigm into self-adaptive mechanism, rejecting 94.18% of attacks.
- Low-Latency Detection:** Fed-LSTM-ZT provides 12.4 ms for detection and 3.2 ms for access decision, satisfying the

stringent latency constraints of 6G edge applications.

- Communication-Efficient:** Total communication overhead of 1.08 GB is affordable for 6G backhaul and forms acceptable compromise between collaboration benefits and privacy-preserving goals.

V. CONCLUSION

This study proposed and evaluated an AI-driven Zero Trust cybersecurity framework for 6G networks, which was designed using Federated Learning paradigm, LSTM-based intrusion detection, dynamic trust mechanism, and edge-native algorithms. The developed Fed-LSTM-ZT framework allowed to scale the threat detection capabilities of individual edge nodes while preserving the privacy of their data. Specifically, the LSTM module enabled intelligent detection of malicious network

activity, which was then used to calculate dynamic trust scores and make access control decisions. The Zero Trust policy engine was responsible for enforcing adaptive authorization rules, including normal access, restricted access, re-authentication, quarantine, and full denial.

The evaluation results demonstrated that the proposed framework achieved high cybersecurity resilience while maintaining network performance. For example, Fed-LSTM-ZT achieved 96.42% intrusion detection accuracy, 95.87% precision, 96.15% recall, 96.01% F1-score, and 0.9845 ROC-AUC. These metrics significantly surpassed those of the centralized LSTM, local LSTM, and rule-based Zero Trust counterparts, thus proving the effectiveness of the privacy-preserving collaborative approach. Moreover, the framework successfully blocked 94.18% of all attacks with only 12.4ms of detection latency and 3.2ms of decision latency. Thus, it met the stringent performance requirements of 6G networks while providing high-level security. Besides, federated model aggregation improved the generalization ability of the local models, which were trained on different data distributions. While FL introduced extra communication and computational overhead, it prevented data exposure and formed a reasonable balance between collaborative intelligence and data privacy. In conclusion, Fed-LSTM-ZT framework proved conceptually viable for securing the intelligent and heterogeneous networks of future.

REFERENCES

- [1] Buck, C., Olenberger, C., Schweizer, A., Völter, F., & Eymann, T. (2021). Never trust, always verify: A multivocal literature review on current knowledge and research gaps of zero-trust. *Computers & Security*, 110, 102436. doi:10.1016/j.cose.2021.102436.
- [2] Fu, Y., Xie, Y., Yi, W., Poudel, B., Cao, J., & Li, H. (2025). IPO-ZTA: An intelligent policy orchestration Zero Trust architecture for B5G and 6G. *Computer Networks*, 269, 111450. doi:10.1016/j.comnet.2025.111450.
- [3] Giordani, M., Polese, M., Mezzavilla, M., Rangan, S., & Zorzi, M. (2020). Toward 6G networks: Use cases and technologies. *IEEE Communications Magazine*, 58(3), 55–61. doi:10.1109/MCOM.001.1900411.
- [4] Hnamte, V., Nguyen, H.-N., Hussain, J., & Kim, Y. H. (2023). A novel two-stage deep learning model for network intrusion detection: LSTM-AE. *IEEE Access*, 11, 37131–37148. doi:10.1109/ACCESS.2023.3266979.
- [5] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60. doi:10.1109/MSP.2020.2975749.
- [6] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research*, 54, 1273–1282.
- [7] Nguyen, V.-L., Lin, P.-C., Cheng, B.-C., Hwang, R.-H., & Lin, Y.-D. (2021). Security and privacy for 6G: A survey on prospective technologies and challenges. *IEEE Communications Surveys & Tutorials*, 23(4), 2384–2428. doi:10.1109/COMST.2021.3108618.
- [8] Niknam, S., Dhillon, H. S., & Reed, J. H. (2020). Federated learning for wireless communications: Motivation, opportunities, and challenges. *IEEE Communications Magazine*, 58(6), 46–51. doi:10.1109/MCOM.001.1900461.
- [9] Park, J., Samarakoon, S., Bennis, M., & Debbah, M. (2019). Wireless network intelligence at the edge. *Proceedings of the IEEE*, 107(11), 2204–2239. doi:10.1109/JPROC.2019.2941458.
- [10] Porambage, P., Gür, G., Moya Osorio, D. P., Liyanage, M., Gurtov, A., & Ylianttila, M. (2021). The roadmap to 6G security and privacy. *IEEE Open Journal of the Communications Society*, 2, 1094–1122. doi:10.1109/OJCOMS.2021.3078081.
- [11] Saad, W., Bennis, M., & Chen, M. (2020). A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. *IEEE Network*, 34(3), 134–142. doi:10.1109/MNET.001.1900287.
- [12] Tataria, H., Shafi, M., Molisch, A. F., Dohler, M., Sjöland, H., & Tufvesson, F. (2021). 6G wireless systems: Vision, requirements, challenges, insights, and opportunities. *Proceedings of the IEEE*, 109(7), 1166–1199. doi:10.1109/JPROC.2021.3061701.
- [13] Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 5, 21954–21961. doi:10.1109/ACCESS.2017.2762418.
- [14] Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE*, 107(8), 1738–1762. doi:10.1109/JPROC.2019.2918951.

Citation of this Article:

Omorogie Michael, Chekwube Georgina Nwankwo, Osita Miracle Nwakeze, & Agubosim Chuka Charles. (2026). Design and Development of an AI-Driven Zero Trust Cybersecurity Framework for 6G Networks Using Federated Learning. *Current Journal of Engineering and Science Research*. 3(9), 30-42. Article DOI: <https://doi.org/10.47001/CJESR/2026.309003>

***** End of the Article *****