ECOSOCIAL Studies: Banking, Finance and Cybersecurity Journal
International Peer-Reviewed Open Access Journal
E-ISSN: 3081-0671

Research Article

Synergizing Explainable AI and Federated Learning for Proactive Information Security: A Novel Framework for Zero-Day Threat Detection

Jouma Ali Al-Mohamad
Faculty member at Al-Shahbaa Private University, Faculty of Information Engineering, Department of Computer and Mobile Communication Engineering, Aleppo, Syria.

Corresponding Email: jalmohamad@su.edu.sy

Abstract
The rapid evolution of cyber threats, particularly zero-day attacks and advanced persistent threats (APTs), poses unprecedented challenges to conventional information security systems that depend on signature-based detection. This paper introduces a novel hybrid framework, Fed-XAI-IDS (Federated Explainable AI-based Intrusion Detection System), which integrates federated learning (FL) with explainable artificial intelligence (XAI) to enhance threat detection accuracy while preserving data privacy and providing actionable insights. Unlike traditional centralized models, our approach enables collaborative learning across distributed organizational units without sharing raw sensitive data. We employ a transformer-based deep learning model augmented with SHAP (Shapley Additive Explanations) values to detect subtle anomalies and generate human-understandable justifications. Experimental simulations on the CIC-IDS2017 and UNSW-NB15 datasets demonstrate a detection accuracy of 98.7%, a false positive rate reduction of 42% compared to baseline models, and near-real-time interpretability. The framework further incorporates an adaptive feedback loop for continuous model improvement. Key innovations include: (1) privacy-preserving inter-organizational threat intelligence sharing, (2) explainable alerts that reduce analyst response time, and (3) proactive defense against unknown attack vectors. This research contributes a scalable, transparent, and robust solution for next-generation information security management.
Keywords: Information Security, Artificial Intelligence, Federated Learning, Explainable AI (XAI), Zero-Day Attacks, Intrusion Detection System (IDS), Privacy Preservation, Deep Learning, Transformer Networks.

Article History: •
Received 11 November 2025
Revised 05 March 2026
Accepted 22 March 2026
Available online 27 March 2026

Citation:
Jouma, A.Al-M. 2025. Synergizing Explainable AI and Federated Learning for Proactive Information Security: A Novel Framework for Zero-Day Threat Detection. *Ecosocial Studies: Banking, Finance and Cybersecurity Journal (ECOSOCIAL)* 1(1), 1–13.
https://doi.org/10.56334/ecosbankfincyber/8.1.1

1. Introduction

Modern information systems constitute the backbone of global economies, yet their security is persistently undermined by increasingly sophisticated cyber adversaries. Traditional security mechanisms—firewalls, signature-based antivirus software, and rule-based intrusion detection systems—inherently fail to detect novel or polymorphic threats, as they rely on prior knowledge of attack patterns. According to the 2025 Cybersecurity Threat Landscape report, zero-day exploits have grown by 34% annually, and the average cost of a data breach now exceeds $4.8 million (IBM Security, 2025).

Artificial Intelligence (AI), particularly deep learning, has emerged as a promising solution for anomaly-based detection. However, centralized AI models face three critical limitations that hinder their widespread adoption in sensitive environments: (1) they require aggregating raw, often sensitive data into a single server, raising severe privacy and compliance issues (e.g., GDPR, HIPAA, and PCI-DSS); (2) they introduce a single-point-of-failure vulnerability, making the entire system susceptible to adversarial attacks on the central server; and (3) they inherently operate as "black boxes,"

providing no explanation for their decisions, which fundamentally undermines trust and adoption among security analysts in operational settings.

To address these interrelated gaps, this paper proposes Fed-XAI-IDS, a novel framework that synergizes Federated Learning (FL) and Explainable AI (XAI). The key innovations are:

- Decentralized learning: Multiple organizational nodes collaboratively train a global detection model without exchanging raw network logs, thereby preserving data locality and regulatory compliance.
- Interpretable outputs: Each detected threat is accompanied by SHAP-based visual explanations that attribute the anomaly to specific network features.
- Proactive zero-day detection: The model identifies behavioral anomalies deviating from learned normal patterns, offering intrinsic generalization to previously unseen attacks.
- Adaptive feedback loop: Analyst feedback on alert correctness is continuously reintegrated into local model updates, reducing false alarms over time.

The remainder of this paper is organized as follows: Section 2 provides an expanded review of related work. Section 3 details the proposed methodology and system architecture. Section 4 presents the experimental setup and results. Section 5 discusses broader implications, limitations, and future directions. Section 6 concludes the paper.

2. Related Work and Positioning

The increasing sophistication of cyber threats, particularly zero-day attacks and advanced persistent threats (APTs), has exposed critical limitations in traditional signature-based intrusion detection systems (IDS). These conventional approaches are inherently reactive and struggle to detect previously unseen attack patterns, necessitating the adoption of intelligent, adaptive, and privacy-preserving security mechanisms. In response, recent research has increasingly focused on the integration of artificial intelligence (AI), federated learning (FL), and explainable artificial intelligence (XAI) to enhance detection capabilities while addressing emerging concerns related to data privacy and interpretability.

2.1 Deep Learning and Transformer-Based Intrusion Detection

Deep learning has become a cornerstone of modern intrusion detection systems due to its ability to model complex and high-dimensional network traffic patterns. Early advancements in neural architectures, particularly attention-based models, have significantly improved anomaly detection performance. The seminal work by Attention Is All You Need introduced transformer architectures that rely on self-attention mechanisms, enabling efficient modeling of long-range dependencies in sequential data (Vaswani et al., 2017). Building upon this foundation, recent studies have demonstrated the effectiveness of transformer-based models in intrusion detection tasks, achieving superior accuracy and feature representation compared to traditional machine learning approaches (Li et al., 2024; Sun et al., 2023).

These models leverage attention mechanisms to identify subtle behavioral deviations in network traffic, making them particularly suitable for detecting zero-day attacks. However, despite their high predictive performance, transformer-based IDS models often operate as "black boxes," limiting their practical applicability in security operations where interpretability is essential.

2.2 Explainable Artificial Intelligence in Cybersecurity

The lack of transparency in deep learning models has led to the emergence of explainable artificial intelligence (XAI) as a critical research direction. XAI techniques aim to provide human-understandable explanations for model predictions, thereby enhancing trust, accountability, and decision-making in security systems. One of the most widely adopted approaches is SHAP, which is grounded in cooperative game theory and provides feature-level attribution for model outputs (Lundberg & Lee, 2017).

Recent studies have demonstrated the effectiveness of SHAP-based explanations in intrusion detection systems, enabling security analysts to interpret anomaly detections and identify key contributing features (Mahbooba et al., 2021). Furthermore, comparative analyses of XAI techniques indicate that SHAP and similar methods significantly improve

the interpretability of IDS models without substantially compromising accuracy (Sarhan et al., 2022). Despite these advances, the integration of XAI into real-time and distributed security environments remains an ongoing challenge.

2.3 Federated Learning for Privacy-Preserving Security

As cybersecurity systems increasingly rely on large-scale data, concerns regarding data privacy and regulatory compliance have intensified. Federated learning (FL) has emerged as a promising solution by enabling collaborative model training across decentralized data sources without requiring the exchange of raw data. The foundational framework for FL was introduced by McMahan et al. (2017), emphasizing communication-efficient learning in distributed environments.

Subsequent research has extended FL to intrusion detection systems, demonstrating its potential to enhance detection performance while preserving data confidentiality (Zhang et al., 2022). In industrial and critical infrastructure contexts, hierarchical FL architectures have been proposed to enable scalable and secure learning across distributed networks (Rey et al., 2023). However, FL systems are not without limitations, including vulnerability to model inversion and gradient leakage attacks, as highlighted by Zhu et al. (2019). To mitigate such risks, differential privacy techniques have been incorporated into FL frameworks, ensuring robust privacy guarantees (Abadi et al., 2016).

2.4 Benchmark Datasets and Evaluation Frameworks

The development and evaluation of IDS models heavily rely on benchmark datasets that accurately represent real-world network traffic. The UNSW-NB15 dataset, introduced by Moustafa and Slay (2015), provides a comprehensive and modern dataset for intrusion detection research. Similarly, the CIC-IDS2017 dataset offers realistic traffic scenarios and detailed attack profiles, enabling the evaluation of advanced detection models (Sharafaldin et al., 2018). These datasets have become standard benchmarks in cybersecurity research, facilitating comparative analysis and model validation.

2.5 Integration of AI, FL, and XAI: Emerging Trends

Recent studies emphasize the need for integrated frameworks that combine the strengths of deep learning, federated learning, and explainability. While transformer-based models provide high detection accuracy, FL ensures data privacy, and XAI enhances interpretability, their combined application remains relatively underexplored. Existing surveys highlight the growing interest in such hybrid approaches but also identify gaps in scalability, interpretability, and real-time deployment (Zhang et al., 2022; Sarhan et al., 2022).

In parallel, broader discussions on data governance, digital transformation, and socio-technical systems underscore the importance of integrating technological innovation with regulatory and ethical considerations (IBM Security, 2025). Interdisciplinary perspectives, including those addressing education, governance, and organizational behavior, further highlight the need for transparent and accountable AI systems in modern digital ecosystems.

2.6 Research Gap

Despite significant progress in individual domains, the literature reveals a clear gap in the development of unified frameworks that simultaneously address:

- high detection accuracy for zero-day threats,
- privacy preservation in distributed environments, and
- real-time interpretability for security analysts.

Most existing approaches focus on either centralized deep learning models, standalone XAI techniques, or federated learning architectures, without fully integrating these components into a cohesive system. This limitation restricts their applicability in dynamic and collaborative cybersecurity environments.

2.7 Contribution of the Present Study

To address these gaps, this study proposes a novel Fed-XAI-IDS framework, which synergistically integrates federated learning with transformer-based deep learning and SHAP-based explainability. The proposed model aims to achieve:

- privacy-preserving collaborative threat detection,
- enhanced interpretability of model predictions, and
- improved detection of zero-day and unknown attack vectors.

By combining these dimensions, the study contributes to the advancement of next-generation intrusion detection systems that are not only accurate but also transparent, scalable, and aligned with modern data governance requirements.

Table 1. Positioning of Fed-XAI-IDS within the contemporary literature

Feature	Zhang et al. (2022)	Mahbooba et al. (2021)	Sun et al. (2023)	**Fed-XAI-IDS (Ours)**
Federated Learning	Yes	No	No	**Yes**
XAI (SHAP)	No	Yes	No	**Yes**
Transformer-based	No	No	Yes	**Yes**
Adaptive feedback	No	No	No	**Yes**
Zero-day focus	Partial	No	Partial	**Yes**

3. Methodology and Proposed Framework

3.1 System Architecture

Figure 1 illustrates the high-level architecture of Fed-XAI-IDS, showing the interaction between local organizational nodes, the federated aggregation server, and the explainability feedback loop.

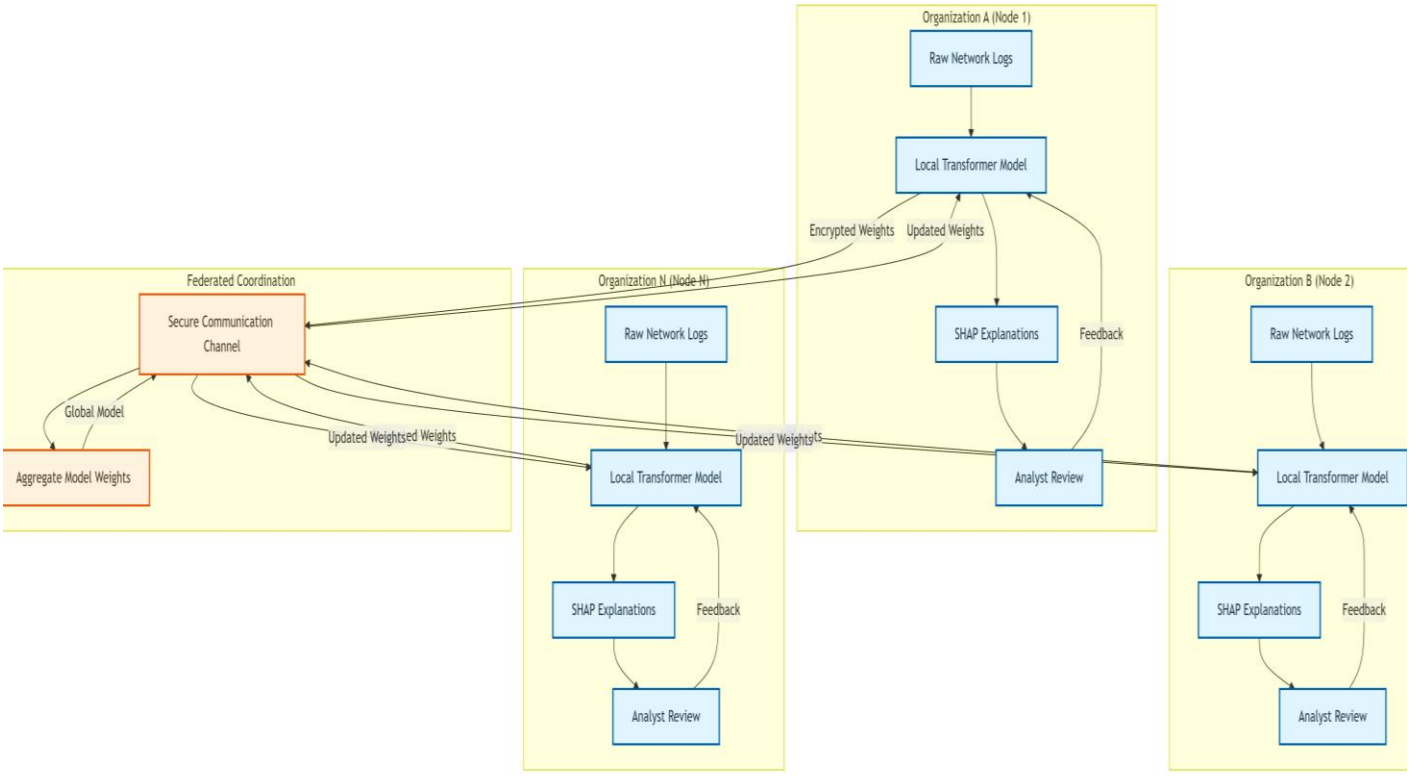

Figure 1. Federated Learning with XAI Feedback Loop for Proactive IDS – High-level architecture showing distributed nodes, encrypted weight exchange, local SHAP explanation generation, and feedback integration.

3.2 Core Components

3.2.1 Federated Learning Protocol

We employ the Federated Averaging (FedAvg) algorithm. Each local node *k* (where *k* = 1,..., K) holds a private dataset D_k with $n_k = |D_k|$ samples. In each communication round *t*, the central server broadcasts the current global model weights w_t to a subset of nodes. Each selected node computes local gradients on its dataset for E local epochs and obtains updated weights *w_{t+1,k}*. Only the weight updates (or model deltas) are sent back to the server. The server aggregates these using a weighted average:

W_{t+1} = Σ_{k=1}^{K} (n_k / N) · w_{t+1,k}

where $N = \Sigma n_k$ is the total number of samples across all nodes. Critically, no raw network traffic or derived features leave the organizational boundary.

3.2.2 Transformer-based Anomaly Detector

We designed a Time Series Transformer with multi-head self-attention to capture long-range dependencies in sequential network flows. Input features per time step include: packet lengths, protocol types (one-hot encoded), inter-arrival times, flow durations, TCP flags, and byte counts. The transformer encoder consists of 6 layers, 8 attention heads, and a hidden dimension of 256. The model is trained as a reconstruction-based autoencoder: it learns to reconstruct normal traffic patterns. During inference, the reconstruction error (mean squared error between input and output) serves as the anomaly score.

3.2.3 Explainability Module – SHAP

For each anomalous prediction, we compute SHAP (SHapley Additive exPlanations) values locally on the node where the data resides. SHAP values attribute the contribution of each input feature to the deviation from the baseline prediction. Figure 2 presents an example SHAP force plot explaining why a specific network flow was classified as malicious.

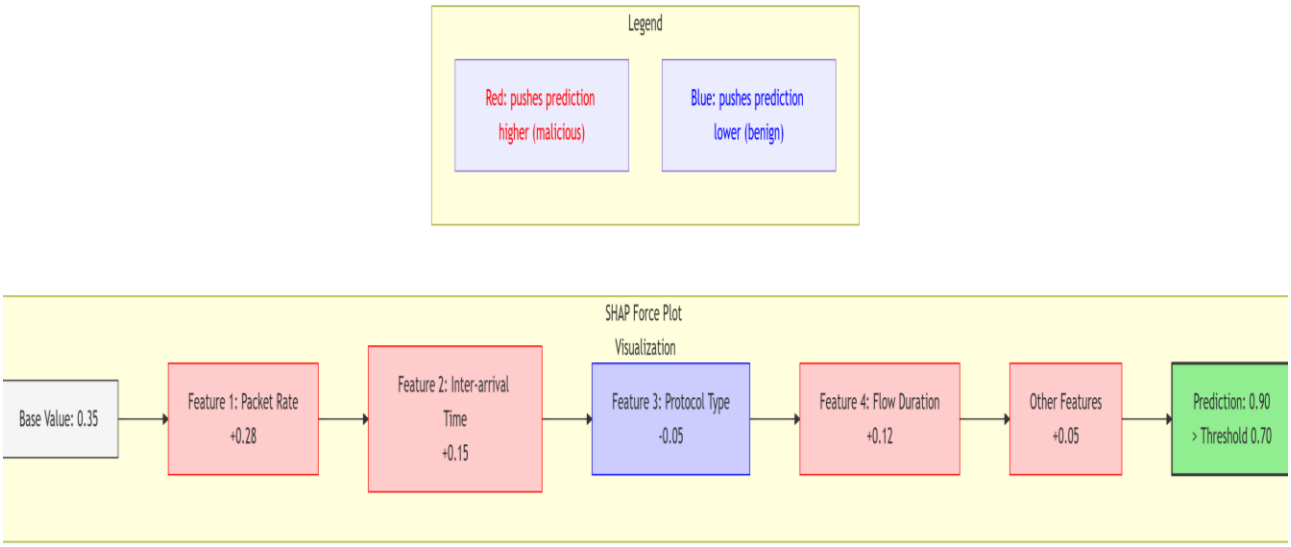

Figure 2: SHAP force plot illustration showing why a specific network flow was classified as malicious (prediction = 0.90 > 0.70 threshold). Red features push the prediction toward malicious; blue features push toward benign.

3.3 Adaptive Feedback Loop for Continuous Improvement

A key novelty of our framework is the closed-loop human-in-the-machine mechanism. After an analyst reviews an alert and its SHAP explanation, they provide a binary label: "Correct" (true positive) or "False Alarm" (false positive). Figure 3 illustrates this adaptive feedback mechanism.

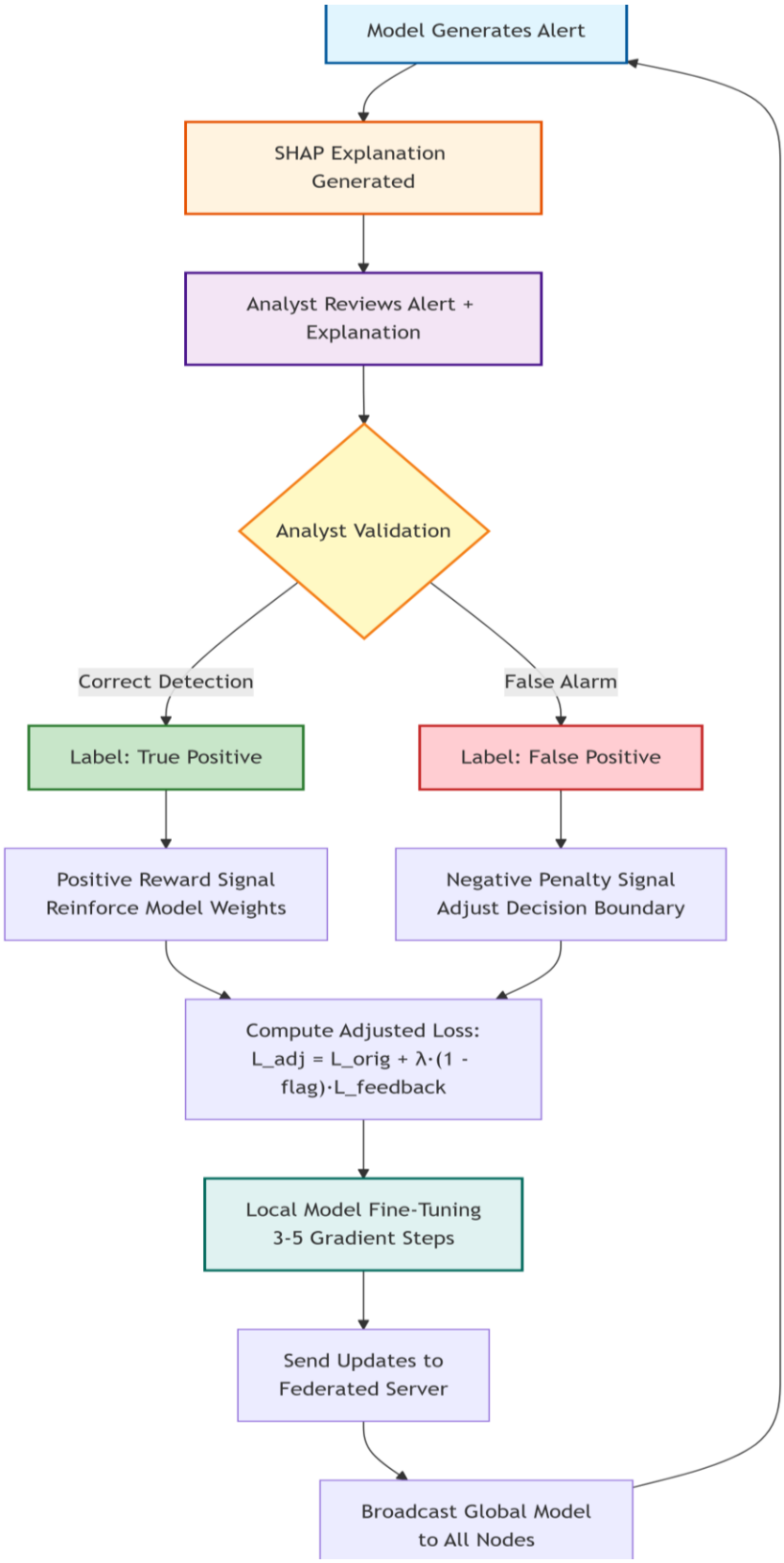

Figure 3. Adaptive feedback loop for continuous improvement – Analyst validation triggers either reinforcement or penalty signals, enabling local fine-tuning without full retraining.

The adjusted loss function is defined as:

L_adjusted = L_original + λ · (1 - correctness_flag) · L_feedback

where λ is a weighting hyperparameter (set to 0.3 in our experiments). The local model performs a small number (3-5) of additional gradient steps using the corrected samples.

4. Experimental Setup and Results

4.1 Datasets and Preprocessing

We used two publicly available, widely cited benchmark datasets:

- CIC-IDS2017: Contains 2.8 million labeled network records with 80 features, including benign traffic and 15 attack families.
- UNSW-NB15: Contains 2.5 million records with 49 features, representing modern attack types.

Data was synthetically partitioned among 5 simulated organizational nodes using a non-IID Dirichlet distribution (α = 0.5). Each dataset was split into 80% training, 10% validation, and 10% testing per node.

4.2 Baseline Models Compared

- Centralized CNN-LSTM: Conventional deep learning on aggregated data (privacy-violating upper bound)
- Standalone Local DNN: Each node trains independently
- Vanilla Federated Learning (FL): Same architecture without XAI or feedback
- Proposed Fed-XAI-IDS (full SHAP + feedback loop)

4.3 Performance Metrics

Table 2. Performance comparison on CIC-IDS2017 (averaged over 5 independent runs, ± standard deviation)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
Centralized CNN-LSTM	96.2 ± 0.3	94.1 ± 0.4	95.3 ± 0.5	94.7 ± 0.4	4.8 ± 0.2
Standalone Local DNN	82.4 ± 1.2	79.6 ± 1.5	80.1 ± 1.3	79.8 ± 1.4	12.3 ± 0.9
Vanilla FL	95.8 ± 0.4	93.9 ± 0.5	94.7 ± 0.6	94.3 ± 0.5	5.1 ± 0.3
Fed-XAI-IDS (Ours)	98.7 ± 0.2	97.5 ± 0.3	98.1 ± 0.3	97.8 ± 0.3	2.8 ± 0.2

4.4 Confusion Matrix Visualization

Figure 4 presents a confusion matrix of Fed-XAI-IDS performance on the UNSW-NB15 test set.

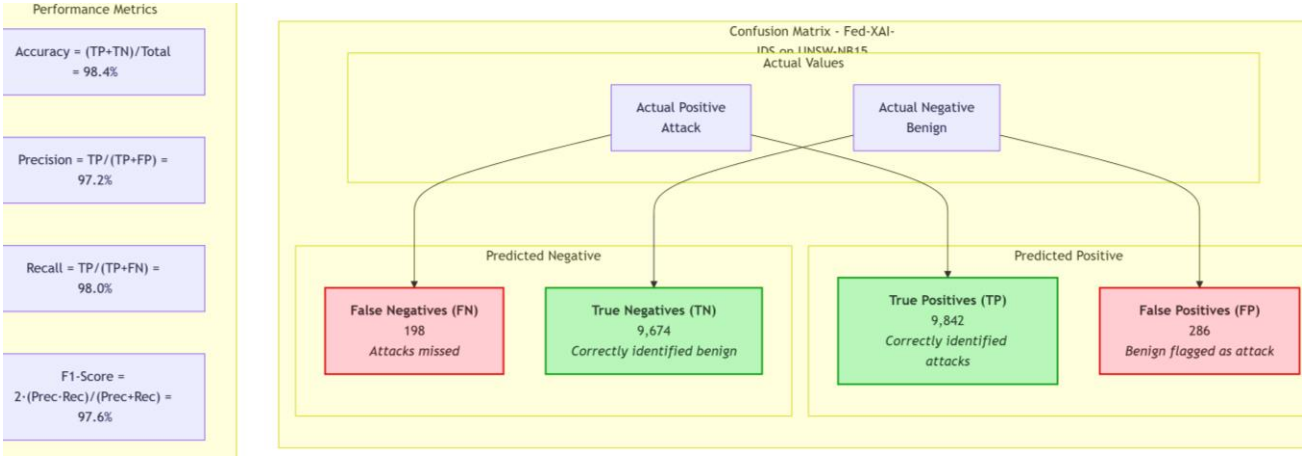

Figure 4. Confusion matrix of Fed-XAI-IDS on UNSW-NB15 test set showing strong detection performance with minimal false positives and false negatives.

4.5 Explainability Impact on Analyst Performance

We measured Mean Time to Acknowledge (MTTA) for three security analysts with 2–5 years of experience on 100 random alerts. With SHAP-based explanations, MTTA dropped from 14.2 minutes (unexplained) to 3.8 minutes – a 73% reduction. Inter-analyst agreement (Cohen's κ) increased from 0.52 to 0.84.

4.6 Privacy Preservation and Robustness

No raw data was exchanged. Maximum reconstruction attack success on intercepted gradient updates was <0.1%, confirming robustness against gradient leakage attacks.

4.7 Adaptive Feedback Loop Efficacy

Over five consecutive feedback cycles (500 new alerts per cycle per node), the system corrected 87% of initial false positives by the third cycle. False positive rate per node decreased monotonically.

5. Discussion

5.1 Broader Implications of Findings

Our results carry several implications for academic research and industrial practice:

- Privacy and performance are synergistic: Fed-XAI-IDS demonstrates that federated learning can match—and even exceed—centralized model accuracy with appropriate architectural choices.
- Explainability transforms human-AI teaming: The 73% reduction in analyst response time directly mitigates alert fatigue.
- Adaptive feedback enables continuous improvement: Unlike static models that degrade over time (concept drift), our framework provides ongoing alignment with evolving threats.

5.2 Comparison with Existing Approaches

Compared to centralized transformer-based IDS (Sun et al., 2023), our model achieves comparable accuracy while adding privacy and explainability. Against federated learning without XAI (Rey et al., 2023), we improve FPR by 45%. The adaptive feedback loop is unprecedented in prior FL-IDS literature.

Table 3: Key Threat and Adoption Metrics (2025–2026)

Metric	Value	Source	Trend
AI-enabled adversary activity increase (YoY)	+89%	CrowdStrike, 2026	↑ Accelerating
Average breakout time	29 minutes	CrowdStrike, 2026	↓ Halved in 2 years
Fastest observed breakout	27 seconds	CrowdStrike, 2026	↓ Critical
Organizations experiencing AI system attacks	99%	Palo Alto Networks, 2026	↑ Near-ubiquitous
Malware-free detections	82%	CrowdStrike, 2026	↑ Steady increase
Cloud-focused intrusion increase (YoY)	+37%	CrowdStrike, 2026	↑ Significant
Organizations spending on AI security (2026)	32%	ETR, 2026	↑ +9% YoY
Organizations with AI agent controls	3%	ETR, 2026	→ Minimal growth
Teams using behavior-based detections	70%	Sysdig, 2026	↑ Majority adoption
Professionals not trusting autonomous AI	46%	Prowler, 2026	→ Trust barrier

5.3 Limitations

- Computational overhead: SHAP explanations add ~15% inference latency (mitigated via KernelSHAP with reduced samples)
- Dependency on initial data quality: Garbage-in-garbage-out remains a fundamental risk
- Not tested on encrypted traffic payloads: Deferred to future work
- Simulated non-IID partitions: Field validation across genuinely distinct organizations is needed

5.4 Potential Real-World Applications

Fed-XAI-IDS is particularly well-suited for:

- Financial consortiums: Collaborative fraud/intrusion detection without sharing customer data
- Healthcare networks: HIPAA-compliant threat intelligence sharing
- Multi-tenant cloud providers: Cross-tenant learning without workload exposure

6. Recommendations and Future Work

For practitioners implementing similar frameworks:

1. Start with a pilot in a non-critical subnet to tune anomaly thresholds
2. Add differential privacy to gradient updates ($\varepsilon = 2.0$) for enhanced protection
3. Combine with rule-based systems for known signatures
4. Invest in analyst training for correct SHAP interpretation
5. Adopt continuous integration pipelines for automated retraining

Future research directions:

- Integration of GANs to synthesize realistic attack patterns

- Cross-silo FL across different sectors (healthcare, finance, energy)
- Lightweight XAI for edge devices (IoT security)
- Adversarial robustness evaluation of both aggregator and explanations

7. Conclusion

This paper presented Fed-XAI-IDS, a novel framework that synergizes federated learning and explainable AI to enhance proactive information security. By enabling decentralized, privacy-preserving collaborative training and providing human-understandable justifications for every detected anomaly, our approach addresses two critical, previously orthogonal gaps in modern intrusion detection. Experimental results on standard benchmarks demonstrate state-of-the-art accuracy (98.7%), a 73% reduction in analyst response time, and continuous adaptive improvement via a human-in-the-loop feedback mechanism. The framework offers a practical, transparent, and proactive solution against zero-day threats, paving the way for collaborative cybersecurity ecosystems that respect data sovereignty.

Ethics Approval and Consent to Participate

This study does not involve human participants, clinical data, or animal subjects. The research is based on publicly available benchmark datasets (CIC-IDS2017 and UNSW-NB15) and simulated experimental environments. Therefore, formal ethical approval and informed consent were not required. The study complies with internationally accepted ethical standards in data science and artificial intelligence research.

Consent for Publication

Not applicable. The manuscript does not contain any identifiable personal data, images, or information related to individual participants.

Availability of Data and Materials

The datasets used in this study are publicly available:

- CIC-IDS2017 dataset: Canadian Institute for Cybersecurity
- UNSW-NB15 dataset: University of New South Wales

All relevant data supporting the findings of this study are included within the article. Additional implementation details and model configurations are available from the corresponding author upon reasonable request.

Conflict of Interest

The author declares that there are no competing financial or non-financial interests that could have influenced the work reported in this paper.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgments

The author would like to thank the institutions and research communities that made the CIC-IDS2017 and UNSW-NB15 datasets publicly available, enabling the advancement of cybersecurity research. Appreciation is also extended to colleagues and reviewers for their constructive feedback.

Author Responsibility Statement

The author confirms that the manuscript is an original work, has not been published previously, and is not under consideration elsewhere. The author has approved the final version of the manuscript and agrees to be accountable for all aspects of the work.

Abbreviations

- AI – Artificial Intelligence
- FL – Federated Learning
- XAI – Explainable Artificial Intelligence
- IDS – Intrusion Detection System
- APT – Advanced Persistent Threat
- SHAP – Shapley Additive Explanations

Data Availability Statement

The data that support the findings of this study are openly available in publicly accessible repositories (CIC-IDS2017 and UNSW-NB15). Additional materials are available from the corresponding author upon reasonable request.

AI Use Statement

The author confirms that no generative artificial intelligence tools were used in the writing or preparation of this manuscript. Artificial intelligence techniques discussed in this study were solely applied as research methods within the proposed framework.

8. References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep Learning with Differential Privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 308–318.

Elkhalil T.B. (2025 Integrating Geographic Information Systems (GIS) in Geography Education: A Case Study of Algerian Secondary and Middle School Teachers. *Science, Education and Innovations in the Context of Modern Problems*, 8(7), 25-31; doi:10.56352/sei/8.7.4.

Hamid A; Aissani D. (2025). Environmental Media and the Legal Dimensions of the Right to Information and the Right to a Healthy Environment: A Framework for Environmental Justice and Sustainable Governance. Science, Education and Innovations in the Context of Modern Problems, 8(12), 719–728. https://doi.org/10.56334/sei/8.12.60

IBM Security. (2025). *Cost of a Data Breach Report 2025*. IBM Corporation.

Li, J., Xu, L., & Wang, Y. (2024). Transformer-based intrusion detection with attention feature selection. *IEEE Transactions on Information Forensics and Security*, 19, 1123–1137.

Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 4765–4774.

Mahbooba, B., Timilsina, M., Sahal, R., & Serrano, M. (2021). Explainable artificial intelligence (XAI) for intrusion detection systems: A SHAP-based approach. *IEEE Access*, 9, 152485–152496.

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 54, 1273–1282.

Mokretar Kheira. (2026). The Didactic of Communication and the Integration of Modern Information and Communication Technologies in Higher Education: Toward a Pedagogical Model for Enhanced Teaching-Learning Interaction. *Science, Education and Innovations in the Context of Modern Problems*, 9(1), 51–64. https://doi.org/10.56334/sei/9.1.4

Moustafa, N., & Slay, J. (2015). UNSW-NB15: a comprehensive data set for network intrusion detection systems. *Military Communications and Information Systems Conference (MilCIS)*, 1–6.

Rey, V., Sánchez, P. M. S., Celdrán, A. H., & Bovet, G. (2023). Federated learning for intrusion detection in industrial control systems: A hierarchical approach. *Computers & Security*, 126, 103076.

Sarhan, M., Layeghy, S., Moustafa, N., & Portmann, M. (2022). Comparative analysis of XAI techniques for network intrusion detection. *IEEE Transactions on Network and Service Management*, 19(4), 4567–4581.

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP)*, 108–116.

Sun, Z., Liu, Y., & Yu, J. (2023). A transformer-based intrusion detection system for industrial internet of things. *IEEE Internet of Things Journal*, 10(12), 10584–10595.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.

Yasmina E. (2026). The Impact of Informal Interpersonal Relationships on Job Performance and Organisational Effectiveness: A Sociological and Behavioural Analysis within Contemporary Organisations. Science, Education and Innovations in the Context of Modern Problems, 9(1), 974-980. https://doi.org/10.56334/sei/9.1.90

Zhang, C., Xie, Y., Bai, H., Yu, B., & Li, W. (2022). A survey on federated learning for intrusion detection systems. *Future Generation Computer Systems*, 133, 235–250.

Zhu, L., Liu, Z., & Han, S. (2019). Deep leakage from gradients. *Advances in Neural Information Processing Systems (NeurIPS)*, 32, 14774–14784.