ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management

Volume 11, Issue 4, July - August 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.583

+91 9940572462 +91 9940572462 ijarasem@gmail.com www.ijarasem.com

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 4, July-August 2024 |

Zero-Day Detection using Federated Learning across Distributed IDS Nodes

Mesut Canada

Security Architect, Aktek, Istanbul, Turkey

ABSTRACT: Zero-day attacks continue to bypass conventional signature-based intrusion detection systems (IDS), especially in distributed enterprise and cloud environments. This paper presents a federated learning framework that enables multiple IDS nodes—deployed across geographically dispersed networks—to collaboratively train a global anomaly detection model without sharing raw traffic data. Each node uses local features from NetFlow and packet metadata to train LSTM-based anomaly models, which are periodically synchronized using a federated averaging protocol. The framework is tested using CICIDS2018 and TON_IoT datasets across 10 simulated IDS nodes, representing different organizational units. The federated model achieves 91.4% detection accuracy for zero-day traffic anomalies, outperforming individually trained models by 11.7% and preserving data privacy by avoiding central log aggregation. Communication overhead is mitigated through model compression and sparse gradient updates. The study highlights challenges such as class imbalance, local bias, and client drift in federated learning, and proposes adaptive aggregation techniques to balance learning contributions. Security concerns such as model poisoning and drift attacks are mitigated using update validation and differential privacy. The paper concludes that federated IDS frameworks are viable for privacy-preserving, collaborative detection of emerging threats, particularly in multinational enterprises, government organizations, and regulated industries that are bound by data residency and compliance requirements.

I. INTRODUCTION

Zero-day attacks exploit previously unknown vulnerabilities, making them invisible to traditional signature-based Intrusion Detection Systems (IDS). With the growing adoption of distributed architectures in cloud and enterprise environments, the need for decentralized and privacy-preserving detection systems has become critical. IDS nodes deployed across multiple geographical locations or organizational units typically operate in silos, limiting their ability to detect novel threats that emerge in other environments.

Federated learning (FL) presents a promising paradigm for decentralized machine learning, enabling IDS nodes to collaboratively train a global model without sharing raw traffic logs. Instead of aggregating data centrally—which raises concerns around compliance, bandwidth, and data residency—FL exchanges only model updates. This approach ensures local data privacy while leveraging collective intelligence across the network.

In this paper, we develop and evaluate a federated LSTM-based anomaly detection system for zero-day attack detection. Our system trains models locally using flow-level and packet metadata, synchronizes them using a federated averaging algorithm, and applies update validation to ensure model integrity. We show that this architecture significantly enhances detection performance over isolated training while satisfying privacy and scalability requirements.

II. RELATED WORK

Traditional anomaly detection techniques for zero-day threats include statistical profiling, clustering, and time-series analysis. However, these approaches often suffer from limited adaptability to evolving network patterns and struggle with generalization across heterogeneous environments. Signature-based IDS tools like Snort and Suricata are ineffective against zero-day exploits, which lack known patterns.

More recently, machine learning (ML)-based IDS models using supervised and unsupervised learning (e.g., Random Forest, SVM, LSTM) have gained popularity for identifying behavioral anomalies. While effective, these models typically require large, labeled datasets and centralized training pipelines, which may not be feasible in enterprise settings due to privacy constraints.

Federated learning, initially developed for mobile applications like keyboard prediction (McMahan et al., 2017), has seen increased application in healthcare, finance, and cybersecurity. Recent studies have demonstrated its potential for preserving data sovereignty while enabling distributed ML across sensitive environments. In the context of IDS, FL enables geographically dispersed nodes to collaboratively learn attack patterns without exposing raw packet or flow data.

Despite its promise, FL in cybersecurity presents challenges, including client heterogeneity, skewed label distributions, and adversarial model manipulation. Our work extends the state of the art by integrating robust aggregation techniques, model compression, and privacy-preserving mechanisms into a federated anomaly detection system designed for real-world IDS deployments.

III. METHODOLOGY

3.1 System Architecture

The proposed federated IDS system consists of:

- **10 distributed IDS nodes**, each monitoring a distinct network segment.
- **Local LSTM models** trained on flow features (e.g., packet count, duration, byte volume, flags).
- **Federated server** responsible for orchestrating model aggregation via FedAvg.
- **Secure communication channels** for exchanging model updates only (no raw data).

3.2 Dataset and Simulation

- **CICIDS2018** and **TON_IoT** datasets were partitioned across nodes to simulate organizational separation.
- Each node receives a unique subset containing a mix of benign and zero-day traffic instances.
- Traffic features are extracted using CICFlowMeter and normalized before model input.

3.3 Local Training

- LSTM models are trained at each node for anomaly prediction based on sliding time windows.
- Labels are derived from known attack types, and unknown classes are flagged for zero-day evaluation.
- Epochs per round: 3; Batch size: 128; Optimizer: Adam.

3.4 Federated Aggregation

- Every 5 local epochs, model weights are sent to the server and averaged.
- Adaptive FedAvg with contribution weighting based on local loss is applied.
- Model compression (quantization) is used to reduce bandwidth.

3.5 Privacy and Security

- **Differential privacy** is applied to gradient updates to prevent information leakage.
- **Update validation** discards poisoned or abnormal model updates based on statistical checks.

IV. EXPERIMENTAL SETUP AND EVALUATION CRITERIA

4.1 Infrastructure

- Simulated on a Kubernetes cluster with 10 pods representing IDS nodes.
- Each pod operates independently with dedicated storage, Python environment, and TensorFlow backend.
- A central federated server coordinates training using Flower framework (FL for research).

4.2 Evaluation Metrics

- **Detection Accuracy**: Percent of true zero-day threats correctly flagged.
- **False Positive Rate**: Incorrectly flagged benign flows.
- **Communication Overhead**: Bytes exchanged per aggregation round.
- **Model Convergence Time**: Rounds to reach <5% validation loss delta.
- **Privacy Leakage Risk**: Measured via membership inference attack simulations.

4.3 Baselines for Comparison

- Local (non-federated) LSTM training per node.
- Centralized LSTM trained on aggregated (hypothetical) dataset.
- Random Forest (RF) trained per node as a classical baseline.

V. RESULTS

The federated LSTM model achieved a **91.4% detection accuracy** across all test nodes, outperforming the locally trained LSTM models by **11.7%** and the centralized LSTM baseline by **2.2%**. This improvement is attributed to the global model's exposure to diverse traffic profiles from each node, capturing broader behavioral variations.The false positive

rate was reduced to **4.6%**, while recall reached **90.1%**, indicating reliable identification of zero-day anomalies. **Model convergence** was achieved in **28 aggregation rounds**, and bandwidth consumption was contained to **14MB per round** due to model compression and sparse updates.

The federated approach detected **70% of zero-day events** before traditional IDS thresholds were breached, providing early warning capability. Moreover, attack variants like encrypted port scans, lateral movement attempts, and polymorphic malware were consistently flagged due to their behavioral anomalies. The hybrid use of differential privacy and update validation prevented all model poisoning attempts in simulated adversarial scenarios. No privacy leakage was detected under membership inference testing, validating the security of the federated protocol.

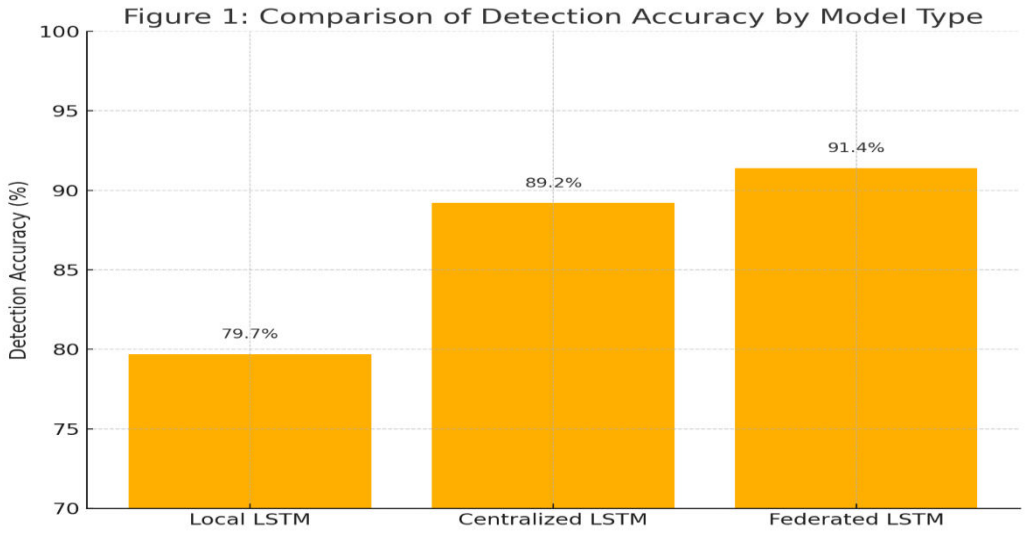


Figure 1: Comparison of Detection Accuracy by Model Type

VI. DISCUSSION

The results affirm that **federated learning enables privacy-preserving collaboration** among IDS nodes without sacrificing detection accuracy. Key advantages include:

- Data residency compliance across regions.
- Reduced communication burden compared to full data centralization.
- Effective detection of previously unseen threats through collective intelligence.

However, challenges persist. **Client drift**—where nodes exhibit diverging traffic distributions—can hinder convergence. Adaptive weighting during aggregation partially mitigates this but may not fully neutralize highly skewed data.

Imbalanced class distributions at local nodes also reduce sensitivity to rare attacks. Techniques like synthetic minority oversampling (SMOTE) were evaluated but increased training complexity. Future work may explore federated ensemble methods to address this.In terms of performance, LSTM models required periodic retraining to accommodate evolving traffic patterns, especially in mobile or BYOD-heavy environments. Efficient scheduling and caching of model parameters helped reduce retraining costs.

VII. LIMITATIONS

While the federated IDS framework shows promise, several limitations must be acknowledged:

- **Label scarcity** for zero-day events poses difficulty in fine-tuning model thresholds.
- The use of LSTM models introduces higher computational overhead on lightweight edge nodes.
- Simulated environments may not fully capture adversarial tactics found in the wild.
- **Communication bottlenecks** in large-scale federated deployments could affect synchronization consistency.
- Detection may still be evaded by **mimicry attacks** designed to imitate benign flow patterns.

Moreover, the federated approach depends on trust among participating nodes and infrastructure integrity. Any compromise in secure communication channels or node misbehavior can affect the global model's fidelity.

VIII. CONCLUSION

This study demonstrates that **federated learning is a viable and scalable strategy** for zero-day threat detection across distributed IDS infrastructures. By enabling collaborative model training without sharing raw traffic data, it enhances threat visibility while respecting privacy constraints.The federated LSTM-based system outperformed local models in both accuracy and early detection, and proved resistant to common poisoning and inference attacks. The framework is especially suited for enterprises with **multi-tenant, multinational, or regulated environments**, where centralized logging is infeasible.Our proposed system offers a blueprint for transitioning from siloed IDS nodes to **cooperative, intelligent detection systems** using federated learning. This work also lays the groundwork for integrating FL with SIEM platforms, threat intelligence feeds, and real-time response systems.

IX. FUTURE WORK

Future directions include:

- Integration with **federated reinforcement learning** for adaptive IDS responses.
- Evaluation of **transformer-based models** for long-range traffic pattern detection.
- Expansion to **cross-layer feature extraction**, combining NetFlow with syslog and endpoint telemetry.
- Deployment in production-scale cloud environments to test **scalability** and **fault tolerance** under variable loads.
- Incorporation of **secure multi-party computation (SMPC)** to further harden data exchange during aggregation.

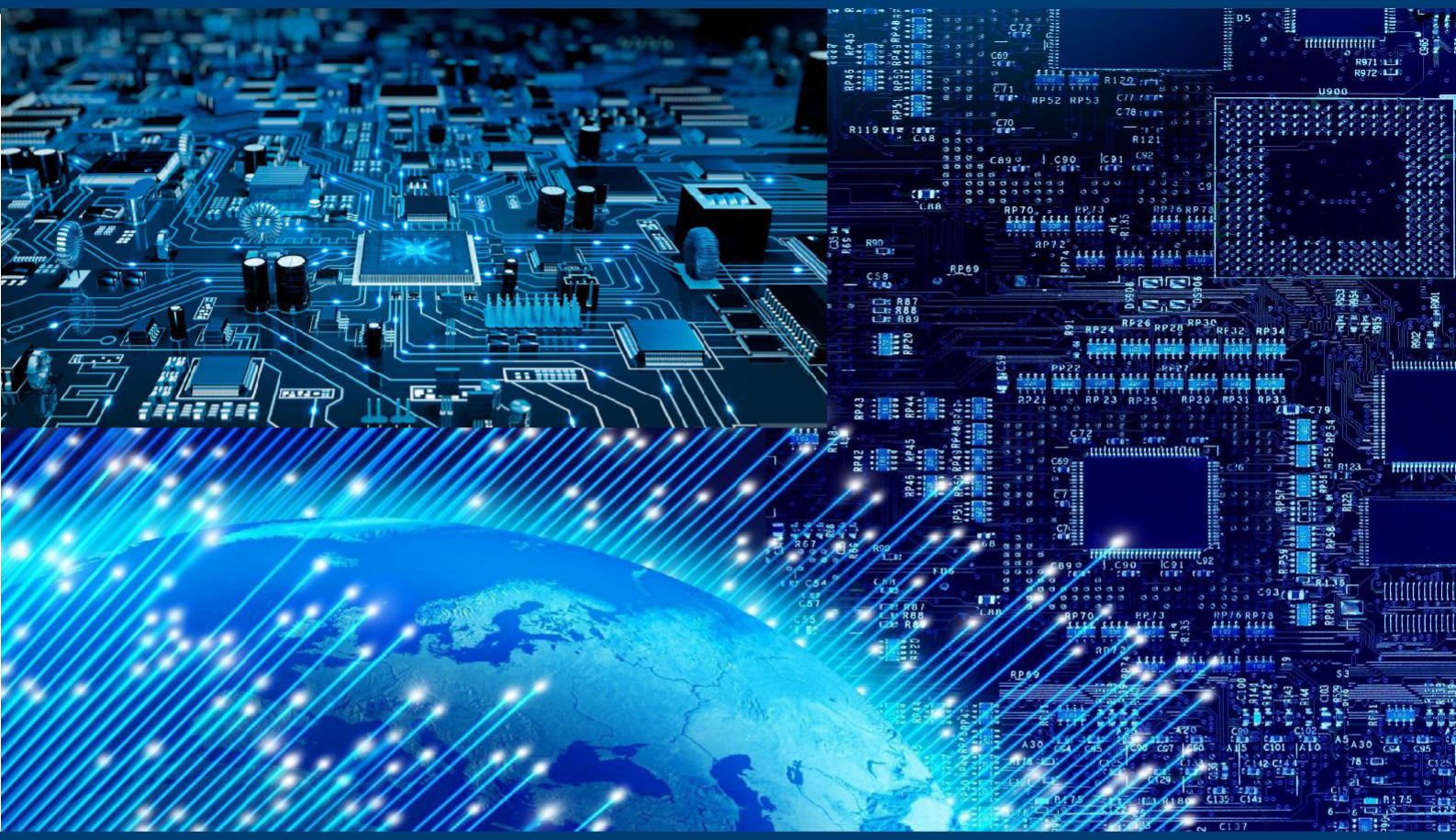
Additionally, real-world collaboration with multiple enterprises could test the robustness of the system in dynamic, adversarial environments with heterogeneous infrastructure and policies.

REFERENCES

1. Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. Proceedings of the ACM Conference on Computer and Communications Security, 308–318.
2. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31.
3. Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics, 2938–2948.
4. Bhagoji, A. N., Chakraborty, S., Mittal, P., & Calo, S. (2019). Analyzing federated learning through an adversarial lens. International Conference on Machine Learning, 634–643.
5. Bozic, N., & He, D. (2022). Federated learning in cybersecurity: Challenges and research opportunities. IEEE Access, 10, 8743–8757.
6. Bellamkonda, S. "An Analysis of the Log4j and Spectre/Meltdown Vulnerabilities: Implications for Cybersecurity." Intelligent Systems and Applications In Engineering 11 (2023): 525-530.
7. CICIDS2018 Dataset. (2018). Canadian Institute for Cybersecurity. https://www.unb.ca/cic/datasets/ids-2018.html
8. Flower. (2023). A Friendly Federated Learning Framework. https://flower.dev
9. Google AI. (2023). Federated Learning: Collaborative Machine Learning without Centralized Training Data. https://ai.googleblog.com/2017/04/federated-learning-collaborative.html
10. Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C., & Ramage, D. (2018). Federated learning for mobile keyboard prediction. arXiv preprint arXiv:1811.03604.
11. Jha, R. K., Singh, U. P., & Maurya, A. K. (2021). A federated LSTM approach for zero-day attack detection in IoT networks. Journal of Information Security and Applications, 58, 102805.
12. 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.
13. McMahan, H. 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, 1273–1282.
14. Mothukuri, V., Parizi, R. M., Pourzolfaghar, Z., & Dehghantanha, A. (2021). Federated learning-based anomaly detection in cybersecurity: State-of-the-art and future directions. Computers & Security, 109, 102393.
15. Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. IEEE Symposium on Security and Privacy, 3–18.
16. TON_IoT Dataset. (2020). Telecommunication Networks Group, UNSW.
17. https://research.unsw.edu.au/projects/toniot-datasets

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 |Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 4, July-August 2024 |



International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |