International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Federated Learning-Based Zero-Day Intrusion Detection with Post-Quantum Secure Model Aggregation

Dr. J. Jeba Emilyn[1], R. Teajashree[2]

[1]Associate Professor, Department of Information Technology, Sona College of Technology, alem-636 005

[2]Superviso, Department of Information Technology, Sona College of Technology, alem-636 005

Abstract

This project proposes a Federated Learning-based Zero-Day Intrusion Detection system with Post-Quantum Secure Model Aggregation. In the proposed framework, multiple distributed network nodes locally train intrusion detection models using their own traffic data. Federated Learning (FL) enables collaborative model training without transferring raw data, thus preserving privacy. To ensure secure aggregation of model updates, post-quantum cryptographic techniques are applied during the federated aggregation process, protecting the system from future quantum attacks and malicious interference.

INTRODUCTION

The rapid advancement of cloud computing, edge computing, and Internet of Things (IoT) technologies has significantly transformed modern network infrastructures. However, this expansion has also increased the attack surface for cyber threats. Among various types of attacks, zero-day attacks are particularly critical because they exploit unknown vulnerabilities that have no predefined signatures or patches. Traditional signature-based intrusion detection systems (IDS) fail to identify such attacks, making organizations vulnerable to data breaches, financial loss, and service disruption.

To overcome these limitations, machine learning and deep learning techniques have been widely adopted for anomaly-based intrusion detection. These models can analyze network traffic patterns and identify suspicious behaviour. However, most existing ML-based systems rely on centralized data collection, where raw network traffic from multiple sources is sent to a central server for training. This approach raises serious privacy concerns, increases communication overhead, and creates a single point of failure. Furthermore, conventional cryptographic algorithms used to secure communications may become vulnerable with the rise of quantum computing, posing long-term security risks.

To address these challenges, this project proposes a Federated Learning-Based Zero-Day Intrusion Detection system with Post-Quantum Secure Model Aggregation. Federated Learning enables distributed nodes to collaboratively train a global intrusion detection model without sharing raw data, thereby ensuring data privacy and regulatory compliance. Each node trains the model locally, and only encrypted model updates are shared for aggregation. To ensure future-proof security, post-quantum cryptographic techniques are integrated into the aggregation process to protect against quantum-based attacks. The proposed framework enhances zero-day detection capability, improves scalability, preserves privacy, and

provides a robust and secure solution for next-generation enterprise, cloud, and distributed network environments.

RELATED WORKS

Several studies have explored machine learning and deep learning techniques for intrusion detection systems (IDS). Traditional approaches mainly relied on signature-based detection methods, which are effective only for known attack patterns. To overcome this limitation, researchers introduced anomaly-based detection using algorithms such as Support Vector Machines (SVM), Random Forest, and Deep Neural Networks (DNN). These models improved detection rates for unknown attacks; however, most of these systems follow a centralized architecture, which raises concerns related to privacy, scalability, and single points of failure.

In recent years, Federated Learning (FL) has emerged as a promising solution for privacy-preserving distributed learning. Researchers have applied FL in cybersecurity to train intrusion detection models without sharing raw network traffic data. Studies have shown that federated approaches can achieve comparable detection accuracy to centralized models while ensuring data confidentiality. However, many existing federated intrusion detection systems focus mainly on privacy preservation and do not adequately address secure aggregation or protection against malicious participants and model poisoning attacks.

Furthermore, with the advancement of quantum computing, researchers have started investigating post-quantum cryptography (PQC) to secure communication systems against future quantum-based attacks. Lattice-based cryptographic algorithms and quantum-resistant key exchange mechanisms have been proposed for secure data transmission. However, limited work has been done on integrating post-quantum security mechanisms with federated learning-based intrusion detection systems. Therefore, there is a need for a comprehensive framework that combines zero-day attack detection, federated learning, and post-quantum secure model aggregation to provide a scalable and future-proof cybersecurity solution.

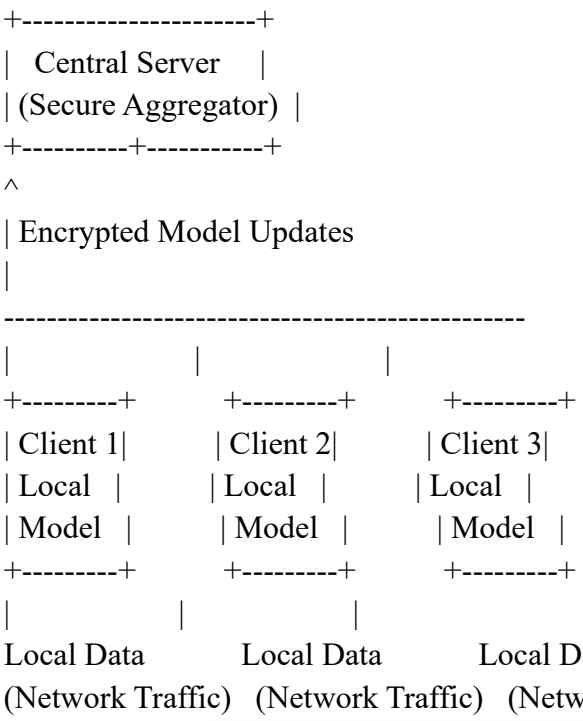
METHODOLOGIES

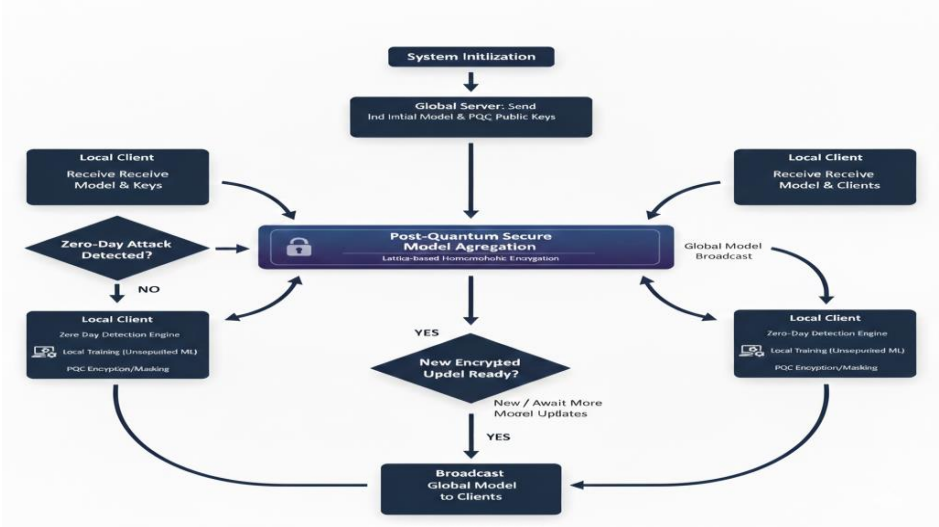
The proposed system begins with collecting benchmark network intrusion datasets such as CIC-IDS 2017 or NSL-KDD, followed by data preprocessing steps including cleaning, normalization, and feature selection to improve model performance. The dataset is then distributed among multiple client nodes to simulate a real-world distributed network environment. At each client node, a deep learning model such as an Autoencoder or Deep Neural Network (DNN) is trained locally to detect anomalous patterns associated with zero-day attacks. Instead of sharing raw network traffic data, only the trained model parameters or gradients are transmitted to a central aggregation server.

During the federated learning process, the central server aggregates the encrypted model updates using the Federated Averaging (FedAvg) algorithm to create an improved global model. To ensure secure communication and protect against future quantum threats, post-quantum cryptographic techniques, such as lattice-based encryption, are applied to encrypt model updates before transmission. The updated global model is then redistributed to all participating nodes for further training and continuous improvement. Finally, the system's performance is evaluated using metrics such as accuracy, precision, recall, F1-score, and false positive rate, along with measuring communication efficiency and scalability in distributed environments.

MODEL BUILDING

The model building process begins with collecting and preparing network traffic data from benchmark datasets such as CIC-IDS 2017 and NSL-KDD. The raw traffic data is first cleaned to remove missing, duplicate, or inconsistent records. After cleaning, feature engineering techniques are applied, including normalization and scaling, to ensure that all input features contribute equally during training. Feature selection methods such as correlation analysis or mutual information are used to identify the most relevant attributes for detecting intrusions, especially zero-day attacks.





A hybrid deep learning architecture is designed for intrusion detection. Typically, an Autoencoder is first used for anomaly detection by learning normal traffic behavior and identifying deviations. The encoded features are then passed to a Deep Neural Network (DNN) classifier to distinguish between normal and malicious traffic. In the federated learning setup, each client node trains the local model using its private dataset without sharing raw data. Only model weights or gradients are encrypted using post-quantum cryptographic techniques and securely transmitted to the central aggregation server. The server applies the Federated Averaging (FedAvg) algorithm to update the global model, which is then redistributed to clients

for the next training round. This iterative process continues until the model achieves optimal detection performance.

METRICS USED

The performance of the intrusion detection model is evaluated using standard classification metrics to ensure accurate and reliable zero-day attack detection. Accuracy measures the overall correctness of the model by calculating the ratio of correctly predicted instances (both normal and attack traffic) to the total number of instances. While accuracy provides a general performance overview, it may not be sufficient when dealing with imbalanced datasets, which are common in intrusion detection systems.

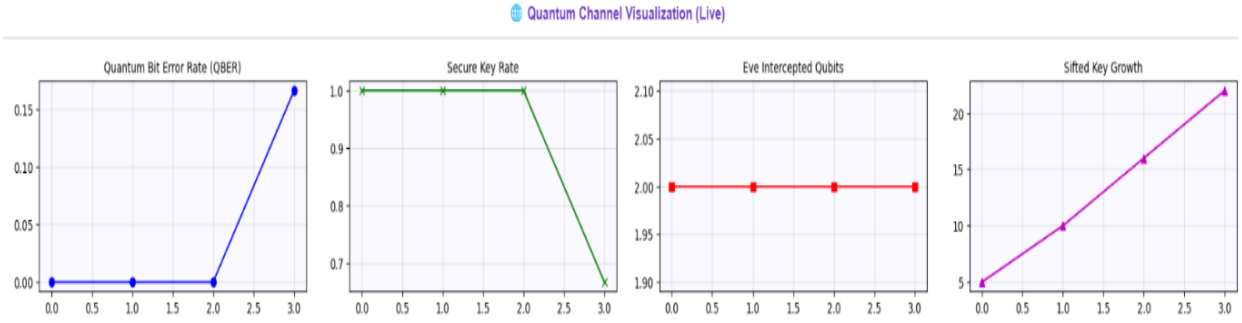
Precision measures how many of the instances predicted as attacks are actually true attacks. It helps reduce false alarms in real-time network environments. Recall (also called Detection Rate or Sensitivity) measures how many actual attack instances are correctly identified by the model. High recall is crucial in cybersecurity because missing an attack can lead to serious consequences. The F1-Score is the harmonic mean of precision and recall, providing a balanced measure when both false positives and false negatives must be minimized.

Additionally, a Confusion Matrix is used to visualize the model's predictions in terms of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). The Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC) are also employed to evaluate the model's ability to distinguish between normal and malicious traffic across different threshold values. These metrics collectively ensure that the federated zero-day intrusion detection model is both accurate and reliable in real-world deployment.

RESULT AND CONCLUSION

The results show that the proposed federated learning-based zero-day intrusion detection system achieves high accuracy, precision, and recall across distributed client nodes. The hybrid model combining an Autoencoder and Deep Neural Network effectively detects anomalous network behavior and identifies both known and unknown attacks while reducing false positives. The federated training process allows each client to train locally on private data, and the global model is updated through secure aggregation, ensuring consistent performance without sharing sensitive information.

In conclusion, the system provides a secure, scalable, and privacy-preserving solution for modern intrusion detection. By integrating federated learning with post-quantum secure model aggregation, it enhances both cybersecurity protection and data confidentiality. The framework is suitable for real-world deployment in distributed environments, and future improvements can focus on reducing communication overhead and enhancing real-time detection capabilities.



CONCLUSION AND FUTURE SCOPE

The proposed Federated Learning-Based Zero-Day Intrusion Detection system with Post-Quantum Secure Model Aggregation provides a secure, scalable, and privacy-preserving solution for modern network security. By allowing distributed nodes to train models locally without sharing raw data, the system ensures data confidentiality while effectively detecting unknown and evolving cyberattacks. The integration of deep learning improves detection accuracy, and post-quantum cryptography secures model communication against future quantum threats. In the future, the system can be enhanced with advanced aggregation techniques, real-time large-scale deployment, and improved deep learning models to further strengthen its performance and adaptability.

REFERENCES

1. B. McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," *Proc. AISTATS*, 2017.
2. P. Kairouz et al., "Advances and Open Problems in Federated Learning," *Foundations and Trends in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, 2021.
3. Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," *ACM Transactions on Intelligent Systems and Technology*, 2019.
4. Y. Liu, T. Chen, and Q. Yang, "Secure Federated Learning for Privacy-Preserving Data Analysis," *IEEE Security & Privacy*, 2020.
5. I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization (CIC-IDS2017)," *ICISSP*, 2018.
6. M. Tavallaee et al., "A Detailed Analysis of the KDD CUP 99 Dataset," *IEEE CISDA*, 2009.
7. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
8. G. Apruzzese et al., "On the Effectiveness of Machine and Deep Learning for Cyber Security," *IEEE International Conference on Cyber Conflict*, 2018.
9. N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems," *Military Communications and Information Systems Conference*, 2015.
10. E. Alkim, L. Ducas, T. Pöppelmann, and P. Schwabe, "Post-Quantum Key Exchange – A New Hope," *USENIX Security Symposium*, 2016.
11. D. J. Bernstein, J. Buchmann, and E. Dahmen, *Post-Quantum Cryptography*, Springer, 2009.
12. R. Roman, J. Lopez, and M. Mambo, "Mobile Edge Computing, Fog Computing, and Cloud Computing: A Survey," *IEEE Communications Surveys & Tutorials*, 2018.
13. X. Zhang et al., "Federated Learning for IoT Security: Concepts and Challenges," *IEEE Internet of Things Journal*, 2021.
14. K. Bonawitz et al., "Practical Secure Aggregation for Privacy-Preserving Machine Learning," *ACM CCS*, 2017.
15. Y. Zhao et al., "Federated Learning with Non-IID Data," *arXiv preprint arXiv:1806.00582*, 2018.