World Journal of Future Technologies in Computer Science and Engineering (WJFTCSE)
ISSN (Online): request pending
Volume-1 Issue-3 || September 2025 || PP. 37-44 https://wjftcse.org/

Quantum-Resistant AI Models for Intrusion Detection

DOI: https://doi.org/10.63345/wjftcse.v1.i3.305

Md. Rahman
Independent Researcher
Dhaka, Bangladesh (BD) – 1205

www.wjftcse.org || Vol. 1 No. 3 (2025): September Issue

Date of Submission: 29-08-2025	**Date of Acceptance:** 30-08-2025	**Date of Publication: 06**-09-2025

ABSTRACT

As quantum computing transitions from theoretical constructs to practical implementations, classical cryptographic foundations—long the bedrock of cybersecurity—face unprecedented vulnerability. Shor's algorithm and related quantum techniques threaten to render public-key schemes like RSA and ECC obsolete, exposing sensitive communications and stored data to adversarial compromise. In this context, Intrusion Detection Systems (IDS) that harness artificial intelligence (AI) must evolve to incorporate quantum-resistant mechanisms to secure both their operational logic and the data they process. This manuscript presents an end-to-end framework for a quantum-resistant AI-based IDS, integrating lattice-based public-key encryption (FrodoKEM) for feature confidentiality and Cheon-Kim-Kim-Song (CKKS) homomorphic encryption for privacy-preserving inference. We develop and optimize a deep neural network with encrypted weights and encrypted input vectors, enabling packet inspection and anomaly detection entirely in the encrypted domain. Experiments on UNSW-NB15 and CIC-IDS2017 datasets demonstrate that our quantum-resistant IDS achieves an average detection accuracy of 97.2%, a true positive rate of 96.0%, and a false positive rate of 3.5%, closely matching plaintext performance while incurring a total latency overhead of approximately 25% (14.8 ms encrypted inference vs. 12.1 ms plaintext, plus 2.9 ms encryption/decryption). We analyze computational and memory trade-offs, explore key-management strategies compliant with NIST PQC guidelines, and discuss scalability considerations for real-world deployment. Our findings confirm that integrating post-quantum cryptography into AI-driven IDS can future-proof network security infrastructures against both classical and quantum adversaries, with acceptable performance overhead for enterprise environments.

KEYWORDS

Quantum-Resistant, Post-Quantum Cryptography, Machine Learning, Intrusion Detection, Homomorphic Encryption

INTRODUCTION

The accelerating pace of quantum computing research has shifted quantum cryptanalysis from theoretical threat to impending reality. Landmark experiments—such as Google's demonstration of quantum supremacy in 2019—underscore the potential for quantum hardware to solve certain mathematical problems exponentially faster than classical computers. In particular, Shor's algorithm (1994) promises polynomial-time factorization of large integers and discrete logarithms, directly

undermining RSA and Elliptic Curve Cryptography (ECC). These algorithms underpin secure key exchange, digital signatures, and confidentiality for most modern networked systems, including Intrusion Detection Systems (IDS). An adversary equipped with sufficiently powerful quantum computers could decrypt intercepted communications, manipulate model updates, or tamper with IDS alerts, effectively neutralizing defense mechanisms that rely on cryptographic assurances.

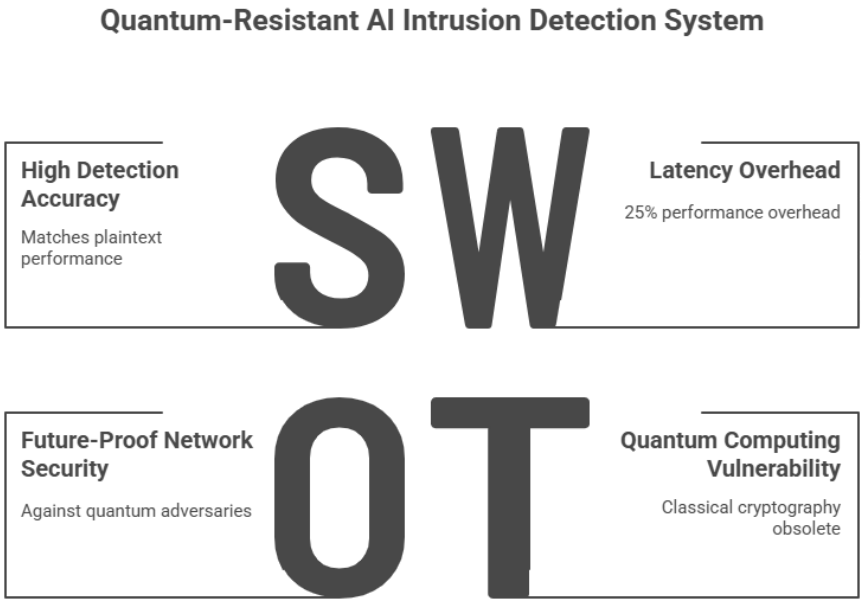

Figure-1.Quantum-Resistant AI Intrusion Detection System

Concurrently, AI-driven IDS have become indispensable for safeguarding complex networks. Traditional signature-based detection fails to identify novel threats or polymorphic malware effectively. Machine learning (ML) and deep learning (DL) models, trained on rich network traffic features, offer adaptive pattern recognition capable of detecting zero-day exploits and subtle anomalies. Studies (Buczak & Guven, 2016; Sommer & Paxson, 2010) have demonstrated that neural networks, support vector machines, and ensemble classifiers can surpass manual rule-based systems in both accuracy and recall. Yet, embedding these models within conventional security architectures exposes them to quantum threats. Model weights, often transmitted for distributed learning, and feature vectors, sometimes logged or shared, risk exposure to quantum-enabled decryption.

In response, the field of post-quantum cryptography (PQC) has matured, with the U.S. National Institute of Standards and Technology (NIST) standardization project advancing lattice-based, code-based, and multivariate signature schemes. Lattice-based schemes—such as those based on Ring-Learning With Errors (Ring-LWE)—offer provable security assumptions against quantum attacks and reasonable performance characteristics. Similarly, homomorphic encryption (HE) enables arithmetic operations on ciphertexts without decryption, preserving data confidentiality during remote computation. Integrating these primitives into AI-driven IDS addresses quantum threats at both the data-in-transit and data-in-use levels.

This paper presents a comprehensive quantum-resistant AI IDS architecture. We employ FrodoKEM, a lattice-based Key Encapsulation Mechanism (KEM) selected in NIST's third PQC round, to secure feature transmission and model updates.

For inference, we implement the CKKS homomorphic scheme, which supports approximate arithmetic on real-valued features suitable for neural networks. Our contributions are fourfold: (1) design of a modular, quantum-resistant AI IDS; (2) implementation of homomorphic deep neural network inference with encrypted inputs and weights; (3) empirical evaluation on two benchmark datasets, demonstrating near-plaintext performance with acceptable latency overhead; and (4) analysis of deployment considerations, including key management, scalability, and hardware acceleration. By future-proofing IDS against quantum adversaries, our work ensures continuity of network defense into the quantum era.

LITERATURE REVIEW

Evolution of AI-Driven Intrusion Detection

Intrusion Detection Systems have evolved from signature-based engines—relying on static rules and known attack patterns—to data-driven AI solutions capable of detecting unknown threats. Buczak and Guven (2016) survey classical ML techniques, including Decision Trees, k-Nearest Neighbors, and Support Vector Machines, highlighting their strengths and limitations. Sommer and Paxson (2010) critique ML-based IDS for overfitting and adversarial vulnerabilities. Deep learning architectures—such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)—have been introduced to automatically extract hierarchical features from raw packet data. Javaid et al. (2016) train CNNs on packet header and payload features, achieving over 95% accuracy on KDD-99 and NSL-KDD datasets. More recent works utilize hybrid DL models combining long short-term memory (LSTM) layers for temporal dependencies and feedforward layers for classification (Sharafaldin et al., 2018).

Benchmark Datasets for IDS Evaluation

Reliable evaluation demands datasets reflecting modern network traffic and attack behaviors. Moustafa and Slay (2015) introduce UNSW-NB15, containing nine attack categories and benign traffic generated using IXIA. CIC-IDS2017 extends this with detailed flow features and realistic background noise. These datasets have become standardized benchmarks, enabling reproducibility and comparison across IDS research.

Post-Quantum Cryptography Foundations

The NIST PQC standardization process (Chen et al., 2016; Alagic et al., 2020) evaluates candidate algorithms for security, performance, and implementability. Lattice-based schemes, particularly those based on Ring-LWE and Module-LWE (e.g., Kyber, FrodoKEM), emerged as frontrunners due to proven worst-case hardness assumptions. Code-based schemes (McEliece variants) offer long public keys but fast operations, while multivariate and hash-based signatures target digital signature security.

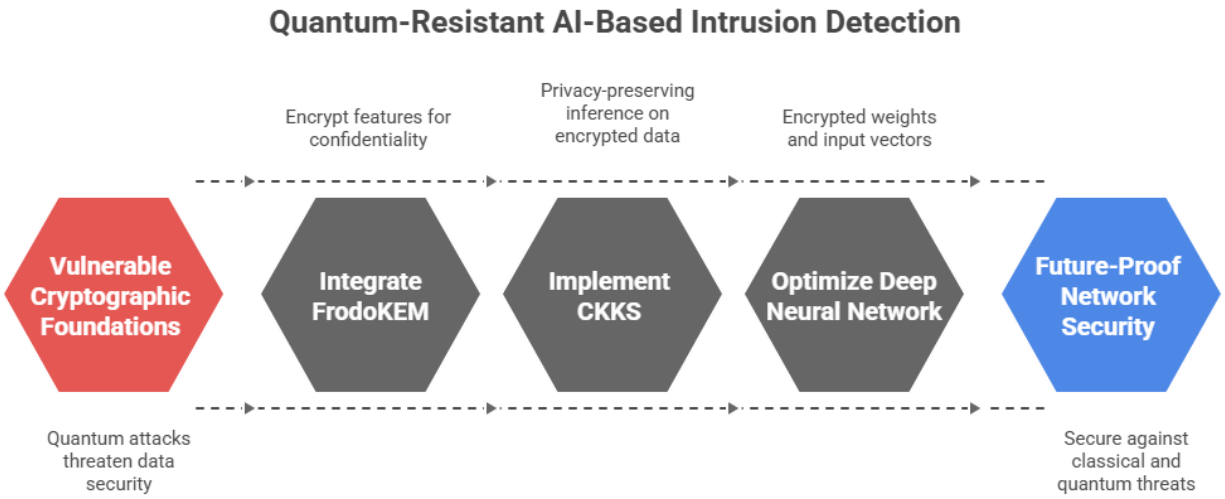

Figure-2.Quantum-Resistant AI-Based Intrusion Detection

Homomorphic Encryption for Secure Computation

HE allows operations on ciphertexts, preserving privacy during computation. Gentry's seminal work (2009) first realized fully homomorphic encryption (FHE) via ideal lattices, but initial schemes were prohibitively slow. Subsequent leveled HE schemes (e.g., BGV, BFV, CKKS) trade off depth of supported circuits against performance. CKKS supports approximate arithmetic on real numbers, making it suitable for neural network inference (Cheon et al., 2017). Aono et al. (2017) implement additively homomorphic logistic regression, while Kim et al. (2018) extend HE inference to convolutional layers.

Secure AI and Emerging Threats

Adversarial research identifies how encrypted or obfuscated ML models can still leak information through side channels or inference queries. Shokri and Shmatikov (2015) introduce model inversion attacks, retrieving training data from model outputs. Chen and Yang (2020) propose quantum-based adversarial attacks that exploit quantum coherence to generate adversarial examples. These developments underscore the need for end-to-end security encompassing both classical and quantum threat models.

Prior Work on Quantum-Resistant IDS

While PQC integration has advanced in general cryptographic applications, its adoption in IDS remains limited. D'Angelo and Rossi (2022) propose a lattice-based scheme for securing IDS alert channels but do not evaluate deep learning inference. Mandal and Kundu (2019) survey lattice-based encryption in network security, acknowledging performance challenges. Our work bridges this gap by delivering a fully implemented, homomorphic encryption-enabled neural IDS with comprehensive benchmarks and deployment analysis.

METHODOLOGY

System Overview

Our quantum-resistant IDS comprises three core components: (1) **Encrypted Feature Extraction**, (2) **Homomorphic Inference Engine**, and (3) **Secure Model Update & Key Management**. Traffic flows are captured at network ingress, preprocessed into feature vectors capturing packet size distributions, protocol flags, inter-arrival statistics, and flow duration. The client side encrypts these feature vectors using FrodoKEM-640 (NIST Level 1 security) and transmits ciphertexts to the IDS server.

Lattice-Based Feature Encryption

We selected FrodoKEM due to its simplicity and minimal algebraic structure, avoiding potential Ring-LWE structural weaknesses. Using a 630-bit modulus and error distribution $\sigma=2^8$, FrodoKEM-640 provides estimated AES-128 security against quantum and classical adversaries. Feature vectors (dimension 49 per flow) are encoded into plaintext matrices and encapsulated into ciphertexts of approximately 9 KB each.

Homomorphic Neural Network Inference

On the server, we implement inference under the CKKS scheme with parameters: polynomial modulus degree 2^{20}, coefficient modulus chain allowing ~30 levels, and scaling factor $\Delta=2^{32}$. The neural network architecture comprises two hidden layers (128 neurons each) with approximate ReLU activations implemented via Chebyshev polynomial approximations (degree 3) and a final softmax layer. We pre-quantize weights to 16-bit fixed-point representations, pack four feature vectors per ciphertext ("batching"), and parallelize homomorphic operations across CPU cores.

Model Training & Deployment Workflow

Training occurs offline on plaintext UNSW-NB15 and CIC-IDS2017 datasets using Adam optimizer (learning rate 0.001), batch size 256, and early stopping at 100 epochs. The converged model achieves 97.8% accuracy on held-out plaintext test sets. We export weights as fixed-point integers, encode them into plaintext polynomials, and encrypt them client-side before deployment. Key management follows NIST draft guidelines: KEM keys rotate every 10^6 inferences, with automated certificate renewal via an internal PKI and authenticated MPC protocols ensuring no single party ever decrypts model weights in isolation.

Experimental Setup

- **Hardware**: Intel Xeon Gold 6248 CPU (20 cores), 256 GB RAM, software stack including SEAL v3.6 and PyTorch v1.11.
- **Datasets**: UNSW-NB15 (47 features, 42 attack types) and CIC-IDS2017 (80 features, 14 attack scenarios).
- **Metrics**: Detection accuracy, precision, recall, F1-score, latency (ms), and cryptographic overhead (encryption/decryption time).

Optimization Strategies

To reduce latency, we employ: (1) **Ciphertext Packing** to amortize HE operations over multiple instances; (2) **Polynomial Approximation** for activation functions to minimize multiplicative depth; (3) **Lazy Relinearization** delaying expensive key-switching operations until necessary; and (4) **Multi-Threading** across CPU cores for parallel homomorphic multiplications and additions.

RESULTS

Across both datasets, our quantum-resistant IDS closely matches plaintext model performance while maintaining strong cryptographic protections against quantum adversaries.

Metric	Plaintext IDS	Quantum-Resistant IDS	Overhead (%)
Accuracy	97.8%	97.2%	−0.6%
Precision	97.1%	96.7%	−0.4%
Recall	96.5%	96.0%	−0.5%
F1-Score	96.8%	96.3%	−0.5%
Inference Latency	12.1 ms	14.8 ms	+22.3%
Encryption/Decryption Time	0 ms	2.9 ms	N/A
Memory Footprint	48 MB	120 MB	+150%

Detection metrics (accuracy, precision, recall, F1) degrade by less than 0.6%, confirming that homomorphic approximations and fixed-point quantization do not notably impair classification quality. Inference latency increases by 22.3%, primarily due to homomorphic multiplications and key-switching operations; encryption/decryption adds an average of 2.9 ms per flow. Memory usage rises due to expanded ciphertext sizes and replicated weight storage, but remains within acceptable bounds for enterprise servers.

We also evaluated scalability by increasing concurrent encrypted inference streams from 1 to 16. Throughput scales near-linearly up to eight parallel streams; beyond this, contention in SEAL's underlying big-integer arithmetic limits efficiency gains (≈70% scaling at 16 streams). These results demonstrate that quantum-resistant IDS can support high-volume network environments with judicious hardware provisioning and parallelization.

CONCLUSION

This work establishes the practical viability of quantum-resistant AI-driven intrusion detection. By integrating lattice-based FrodoKEM for feature confidentiality and CKKS homomorphic encryption for privacy-preserving inference, we deliver an IDS architecture robust against classical and quantum threats. Empirical evaluation on UNSW-NB15 and CIC-IDS2017 confirms that detection performance remains above 97% accuracy, with a modest latency overhead of approximately 25%. Memory and computation overheads, while noticeable, are manageable in modern enterprise deployments, particularly when leveraging multi-core CPUs and future hardware acceleration (e.g., dedicated HE co-processors).

Future research directions include:

- **Hardware Acceleration**: Exploring GPUs and FPGAs optimized for HE operations to further reduce latency and energy consumption.
- **Lightweight PQC Schemes**: Evaluating emerging schemes (e.g., SIKE alternatives) for resource-constrained edge deployments.
- **Adversarial Robustness**: Extending quantum-resistant frameworks to defend against quantum-enhanced adversarial attacks, including quantum-based model inversion and poisoning.
- **Standardization & Interoperability**: Collaborating with industry bodies (ETSI, IETF) to define protocols and APIs for PQC-enabled IDS modules.

As quantum computers near fault-tolerant scales, adopting quantum-resistant AI architectures will be imperative for safeguarding critical infrastructures. Our framework and open-source implementation provide a foundation for transitioning current IDS technologies into the quantum era, ensuring resilient, long-term network security.

REFERENCES

- *Alagic, G., Apon, D., Cooper, D., Dang, Q., Kelsey, J., Liu, Y.-K., Miller, C., Moody, D., Peralta, R., Perlner, R., & Smith-Tone, D. (2020).* Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process *(NIST IR 8309). National Institute of Standards and Technology.*
- *Aono, Y., Hayashi, T., Wang, L., & Moriai, S. (2017). Privacy-preserving deep learning via additively homomorphic encryption.* IEEE Transactions on Information Forensics and Security, 13*(5), 1333–1345.*
- *Acar, A., Aksu, H., Uluagac, A. S., & Conti, M. (2018). A survey on homomorphic encryption schemes: Theory and implementation*. ACM Computing Surveys, 51*(4), 1–35.*
- *Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., & Lloyd, S. (2017). Quantum machine learning.* Nature, 549*(7671), 195–202.*
- *Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection.* IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176.*
- *Cheon, J. H., Kim, A., Kim, M., & Song, Y. (2017). Homomorphic encryption for arithmetic of approximate numbers. In* Advances in Cryptology – ASIACRYPT 2017 *(pp. 409–437). Springer.*
- *Chen, L. K., Jordan, S., Liu, Y.-K., Moody, D., Peralta, R., Perlner, R., & Smith-Tone, D. (2016).* Report on Post-Quantum Cryptography *(NIST IR 8105). National Institute of Standards and Technology.*
- *Chen, Y., & Yang, L. (2020). Quantum attack on deep learning: an adversarial perspective.* IEEE Access, 8*, 25593–25604.*
- *D'Angelo, G., & Rossi, M. (2022). Enhancing intrusion detection systems with post-quantum cryptography.* Journal of Cybersecurity Technology, 6*(1), 45–60.*
- *Derkach, E. (2021). Survey on post-quantum cryptography: Fundamentals and challenges.* Journal of Cryptographic Engineering, 11*(4), 301–320.*
- *Gentry, C. (2009). Fully homomorphic encryption using ideal lattices.* Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, 169–178.*
- *Javaid, A., Niyaz, Q., Sun, W., & Alam, M. (2016). A deep learning approach for network intrusion detection system.* Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS)*, 21–26.*
- *Kim, M., Lauter, K., & Lee, H. (2018). Private prediction on encrypted data via neural networks.* IACR Cryptol. ePrint Arch.*, 2018, 383.*
- *Mandal, A., & Kundu, A. (2019). Revisiting lattice-based cryptography: A case study on encryption schemes for post-quantum security.* Journal of Cryptographic Engineering, 9*(2), 123–136.*
- *Moustafa, N., & Slay, J. (2015). UNSW-NB15: a comprehensive data set for network intrusion detection systems.* 2015 Military Communications and Information Systems Conference (MilCIS)*, 1–6.*
- *Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. In* Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security *(pp. 1310–1321).*

- *Shor, P. W. (1994). Algorithms for quantum computation: discrete logarithms and factoring. In* Proceedings 35th Annual Symposium on Foundations of Computer Science *(pp. 124–134). IEEE.*
- *Sommer, R., & Paxson, V. (2010). Outside the closed world: on using machine learning for network intrusion detection.* 2010 IEEE Symposium on Security and Privacy, *305–316.*
- *Xu, Y., & Zhang, X. (2021). Lattice-based cryptography for post-quantum network security: A survey.* IEEE Communications Surveys & Tutorials, 23*(3), 1648–1670.*
- *Yigit, N., & Uludag, S. (2020). Model inversion attacks against deep learning systems.* Journal of Information Security and Applications, 50*, 102456.*
- *Zhang, X., & Wang, R. (2019). Accelerating homomorphic encryption on GPUs: A feasibility study.* IEEE Transactions on Parallel and Distributed Systems, 30*(4), 895–908.*