INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Forensics Investigation Framework for Advanced Threat Detection in Quantum-Era Networks

Owusu Nyarko-Boateng[1]*, Isaac Kofi Nti[1], Samuel Boateng[1], Adebayo Felix Adekoya[2], Benjamin Asubam Weyori[3], Faiza Umar Bawah[4], Peter Nimbe[4], Foster Yeboah[5], Henrietta Adjei Pokua[6]

1 School of Information Technology, University of Cincinnati, Oh, USA
2 Faculty of Computing, Engineering & Mathematics Science, Catholic University of Ghana
3 Department of Electrical and Computer Engineering, University of Energy and Natural Resources, Ghana
4 Department of Computer Science & Informatics, University of Energy and Natural Resources, Ghana
5 Department of Computer Science, University of Cincinnati, Oh, USA
6 Department of Computer Science, Sunyani Technical University, Sunyani, Ghana

OPEN ACCESS

Received: 11/08/2025
Accepted: 10/11/2025
Published: 07/12/2025

Citation: Nyarko-Boateng O, Nti IK, Boateng S, Adekoya AF, Weyori BA, Bawah FU, Nimbe P, Yeboah F, Pokua HA (2025) Forensics Investigation Framework for Advanced Threat Detection in Quantum-Era Networks. Indian Journal of Science and Technology 18(44): 3524-3543. https://doi.org/10.17485/IJST/v18i44.1401

* **Corresponding author**.

owusu.nyarko-boateng@uenr.edu.gh

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To address the urgent need for forensic systems capable of detecting and analyzing advanced persistent threats in hybrid quantum-classical communication infrastructures, particularly those that may compromise quantum key distribution environments. **Method**: The study introduces a Quantum-Aware Forensics Investigation Framework, a multi-layered forensic architecture combining quantum telemetry, classical metadata analysis, and machine learning-driven threat classification. Experimental validation was conducted using a simulated testbed built with SimulaQron, Wireshark, and custom scripting tools. Various quantum attack scenarios were emulated, including *intercept-resend*, *entanglement flooding*, and *control-plane hijacking*. Machine learning models Random Forest, SVM, and Autoencoder were tested as standalone classifiers. A stacked ensemble model, with Random Forest and SVM as base learners and Logistic Regression as meta-classifier, was implemented for performance optimization. We used an experimentally generated, cross-layer dataset from a SimulaQron BB84 QKD emulation by combining quantum logs and classical control-plane captures under benign and scripted attacks such as intercept–resend, entanglement flooding, payload obfuscation, session hijacking, spoofing. Parameters studied were quantum - QBER, event inter-arrival jitter, event/count rate and classical - packet/flow statistics, inter-arrival mean/variance, latency proxy, TCP SYN/RST flags, byte-level Shannon entropy, with labels for benign vs. attack class. **Findings**: The standalone models achieved moderate performance on the held-out test set for Random Forest: ROC AUC = 0.93, F1 = 0.90, MCC = 0.86, Brier = 0.072; SVM (RBF): ROC AUC = 0.91, F1 = 0.88, MCC = 0.82, Brier = 0.081; Autoencoder (one-class): ROC AUC = 0.87, F1 = 0.83, MCC = 0.74, Brier = 0.094. By contrast, the stacked

ensemble delivered perfect detection metrics for ROC AUC = 1.00, F1 = 1.00, MCC = 1.00, and Brier = 0.014. The study further emphasized the need for forensic systems to support explainability and continuous adaptability via Explainable AI and online learning with drift detection. **Novelty**: This study presents a cross-layer forensic framework for quantum–classical hybrid networks that fuses QKD telemetry with classical control-plane evidence and machine-learning analytics. Unlike prior work that treats these planes separately, our design unifies event-level QKD signals such QBER, arrival-time jitter with packet/flow features to produce timestamp-aligned, explainable alerts. In evaluation, the stacked-ensemble detector achieved perfect detection metrics for ROC AUC, F1, MCC and Brier on held-out data, which distinctly outperformed single-model baselines. The framework couples these gains with an XAI layer and an online, drift-aware learning loop, providing a scalable, auditable, and resilient foundation for forensic intelligence in the quantum era.

Keywords: Quantum network forensics; QKD security; Advanced threat detection; Hybrid quantum-classical networks; Quantum-safe evidence; SimulaQron; Quantum cybersecurity.

1 Introduction

As quantum technologies transition from laboratory prototypes to fielded systems, the assumptions that underpin classical cybersecurity and digital forensics are being re-examined. Quantum networks, which leverage entanglement and quantum key distribution (QKD), promise security grounded in physical law rather than computational hardness. Yet this same physics complicates evidence acquisition and preservation, creating blind spots for classical forensic tooling[1,2]. In parallel, quantum-specific threats e.g., relay manipulation, noise injection, timing disruptions during key exchange extend the attack surface beyond traditional protocol stacks[3]. In real deployments, the challenge is hybrid: classical control/data planes remain susceptible to spoofing and hijacking while quantum links introduce fragile, state-dependent artifacts that are hard to capture with conventional methods[4].

While the security community has invested heavily in post-quantum cryptography (PQC) to protect classical systems against quantum adversaries, PQC does not address forensic logging, cross-layer threat detection, or evidence correlation in quantum communications[5]. Simulation frameworks such as SimulaQron help prototype quantum internetworking but provide no built-in modules for capturing, labeling, and explaining cross-layer evidence under attack[6]. To fill this gap, we introduce the Quantum-Aware Forensics Investigation Framework (QAFIF), a layered, cross-domain approach that fuses quantum telemetry with classical packet/flow analytics, applies machine-learning detectors with calibration, and preserves artifacts via PQC-anchored, tamper-evident logging. On a held-out hybrid dataset, single-model baselines achieve AUC 0.87–0.93; F1 0.83–0.90; MCC 0.74–0.86, whereas a stacked ensemble attains AUC 1.00; F1 1.00; MCC 1.00; Brier 0.014, which demonstrate both high discrimination and strong probabilistic calibration.

1.1 Research Contributions

This study introduces a cross-layer forensic model that formalizes a telemetry schema to timestamp-align quantum signals such as QBER, arrival-time jitter with classical indicators like flow timing, TCP control flags, byte-entropy, producing replayable, evidentiary case files. We quantitatively compare baseline and stacked-ensemble detectors and report explicit calibration (Brier), a dimension under-reported in prior reviews. We further couple detection with explainable, admissible analytics, linking model outputs to XAI artefacts that support legal-grade narratives and chain-of-custody. Finally, we align PQC-anchored logging with NIST FIPS 203/204/205 to strengthen authenticity and non-repudiation[7–10].

1.2 Research Gaps

The reviews conducted on existing literature are (i) security-proof-centric and plane-siloed, (ii) light on operational forensics such as evidence models, explainability, chain-of-custody, (iii) short on benchmark datasets/labels for QKD-centric forensics, and (iv) out-of-date on PQC standards and their logging/authentication implications[7–12].

1.3 Research Questions

1. How can quantum telemetry and classical evidence be fused to detect, explain, and preserve traces of hybrid attacks?
2. What performance and calibration can be achieved by single-model vs. stacked-ensemble detectors on hybrid datasets?
3. Which logging primitives best ensure forensic integrity over time?

4. How should systems incorporate online learning with drift detection to remain reliable under evolving conditions?

1.4 Scope & gaps in prior reviews

Seminal surveys cogently cover QKD security proofs and practice but remain largely siloed from operational forensics, cross-layer telemetry, and explainability needs in real deployments [13]. Newer reviews emphasize QKD networks and key management at scale and catalog practical vulnerabilities/side channels in commercial devices yet typically stop short of end-to-end forensic workflows and data/label standards (attack taxonomies, evidence logging, analyst-facing explanations) [11,14]. Quantum-network testbeds and simulators such as NetSquid, SeQUeNCe, QuNetSim, have matured from discrete-event physical fidelity to full stacks, but only recently expose aligned classical control-plane traces suitable for forensic correlation [15–20]. In cyber-defence, Explainable AI has progressed from concept to comprehensive surveys and toolboxes like SHAP/LIME, yet guidance for legal-grade justifications and chain-of-custody in quantum–classical settings is still nascent [12,21–23]. Finally, NIST's 2024 PQC standards, create new requirements and opportunities for tamper-evident logging and authentication that most prior reviews could not incorporate [7–10,24,25].

1.4.1. Classical Digital Forensics and Intrusion Detection

Classical network forensics relies on packet capture, log analysis, correlation, and anomaly detection. Frameworks such as the OSI-aligned Network Forensics Framework (NFF) enable real-time inspection and trace reconstruction [26]. Tools including Wireshark, NetFlow, and Suricata remain central to attribution workflows. However, these systems assume deterministic traffic features and structured layers that do not transfer directly to quantum communications. IDS have evolved with machine learning, where SVMs, decision trees, and ensembles e.g., Random Forests, deliver high classical threat-classification accuracy [27]; yet these models are built on TCP/IP-centric features absent in quantum protocols.

1.4.2. Security in Quantum Networks

Quantum networks use superposition, entanglement, and measurement disturbance to protect communications; QKD is the flagship application [28]. Despite strong theoretical assurances, practical deployments face side-channel/device vulnerabilities and channel manipulation, including intercept–resend, photon-number splitting, and detector blinding [29,30]. These often manifest as increases in QBER, a signal we later use for forensic detection. Beyond QKD, architectures employing entanglement swapping and teleportation introduces new traceability risks, compromised swapping nodes can disrupt correlations without easy detection [1]. The fragility of quantum states also means conventional "collect-and-inspect" forensics can collapse evidence during observation.

1.4.3. Quantum Forensics: An Emerging Discipline

Literature on quantum forensics is nascent. Early work outlines passive monitoring strategies that infer anomalies from aggregate QBER and channel perturbations but lacks granularity for real-time response. Integrating classical metadata (IP headers, session IDs, time-sync logs) with quantum state transitions has been proposed to improve traceability, aligning with our hybrid approach [31]. While SimulaQron enables controlled experimentation for quantum internet protocols, it lacks forensic trace capture/labeling modules and only limited support for live monitoring of session state and QBER trends, which are crucial for effective forensic capabilities [6].

1.4.4. Post-Quantum Cryptography and Log Integrity

PQC encompasses cryptosystems resistant to quantum attacks (e.g., lattice-, hash-, and multivariate-based) [5]. For forensic readiness, tamper-evident logging and non-repudiation are essential. As blockchain-style audit trails proliferate, integrating PQC signatures such as ML-KEM/ML-DSA/SLH-DSA helps ensure logs remain verifiable against future quantum adversaries [32], with direct alignment to NIST FIPS 203/204/205 practices used in QAFIF.

1.4.5. Machine Learning in Hybrid Network Security

Machine learning models like k-means, autoencoders, RNNs/CNNs supports traffic classification, anomaly detection, and forecasting in hybrid settings. In quantum-aware contexts, ML can track QBER shifts, latency fluctuations, and session mismatches as indicators of compromise. For instance, CNN-based detection of anomalies in QKD signal distributions has been demonstrated in simulation, motivating cross-layer ML designs that we generalize in QAFIF to unify quantum and classical analytics.

1.4.6. Synthesis of Gaps

This review (i) unifies quantum telemetry and classical evidence for cross-layer attribution; (ii) quantifies baseline vs. ensemble detection with calibration; (iii) connects detection to XAI and PQC-anchored logging for non-repudiation; and (iv) maps operational readiness by bridging recent testbeds and demonstrations ML-based detection of QKD device defects/attacks; metro/inter-city links; twin-field and mode-pairing advances, while highlighting missing datasets, labels, and drift-detection needed for reproducible evaluation [14,20,33–35]. Collectively, we position forensics as a first-class design goal for quantum–classical networks, not an afterthought.

This study addresses these gaps by proposing a unified, quantum-aware forensics framework that combines classical packet logging, quantum signal analysis, ML-based detection with calibration, and post-quantum secure evidence preservation.

2 Review Methodology

2.1 Databases and Sources

We surveyed peer-reviewed and standards sources spanning computer science, communications, and quantum engineering: IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect (Elsevier), Nature Portfolio/Communications Physics, Optica/OSA Publishing, AVS Quantum Science, Frontiers, arXiv (preprints) for currency checks, and NIST CSRC for standards and guidance (including FIPS documents). Reference discovery was complemented by backward/forward citation chaining from milestone works already in our list [1–6,26–32,36–39] and the newly added items [7–25,33–35,40].

2.2 Search Strategy

We combined controlled and free-text terms with Boolean operators and field filters. Core blocks:

- Quantum networking & QKD: "quantum network*", "quantum internet", "QKD", "BB84", "twin-field", "mode-pair*", "entanglement swapping", "measurement-device-independent".
- Security & forensics: "forensic*", "incident response", "evidence", "chain-of-custody", "logging", "tamper-evident", "intrusion detection", "anomaly detection".
- Explainability & ML: "explainable AI", "XAI", "feature attribution", "SHAP", "LIME", "ensemble", "calibration", "Brier".
- PQC & standards: "post-quantum cryptograph*", "FIPS 203", "FIPS 204", "FIPS 205", "lattice-based", "hash-based", "ML-KEM", "ML-DSA", "SLH-DSA".
- Simulators/testbeds: "SimulaQron", "NetSquid", "QuNetSim", "testbed", "field trial".

An example of query patterns used in IEEE Xplore is *("quantum key distribution" OR QKD OR "quantum network*") AND (forensic* OR "intrusion detection" OR logging OR "explainable AI" OR XAI) AND (2020:2025).*

2.3 Time Window and Language

Coverage: 2009–2025 to include foundational security and implementation papers, with an a priori target that $\geq$40% of citations be 2020–2025 (met through [7–25,33–35,40]). Only English-language publications were considered.

2.4 Inclusion and Exclusion Criteria

The inclusion criteria used in the search include (i) address QKD or quantum networking with security/operational implications; (ii) contribute to forensics, logging, explainability, datasets, or cross-layer analysis; (iii) provide quantitative results or concrete architectural/standards contributions; (iv) standards and best-practice documents relevant to evidence integrity.

The exclusion criteria used includes: (i) Purely theoretical security-proof papers without operational or measurement implications; (ii) non-reproducible claims with no method details or metrics; (iii) duplicates and non-English items; (iv) tangential works.

2.5 Screening and Selection Workflow

A PRISMA-like process was followed. The initial search identified 1,146 records; 214 duplicates were removed. Titles/abstracts of 932 records were screened; 828 were excluded by criteria above. 104 full texts were assessed; 47 were included in the qualitative synthesis, and a subset informed the quantitative benchmarking and standards synthesis presented in this review. The final reference list was pruned to 40–50 items emphasizing milestone works plus recent (2020–2025) advances, corresponding to [1–40].

2.6 Data Extraction and Appraisal

For each included item we extracted were problem scope, network/protocol context, threat model, dataset availability/labels, metrics, evidence/logging approach, XAI artifacts, and operational constraints. Study quality was appraised as replicability, external validity, and forensic relevance.

3 Overview

To address the forensics challenges introduced by quantum-era communication, we propose the QAFIF. QAFIF is a layered, modular architecture designed to monitor, detect, and analyze anomalies across both classical and quantum communication channels. It bridges the forensic gap between the quantum and classical domains by integrating real-time monitoring, quantum state analytics, classical control-plane logging, and secure evidence preservation mechanisms.

3.1 Architecture Overview

The Quantum-Aware Forensics Investigation Framework is composed of five interconnected components that work together to monitor, analyze, detect, and securely preserve forensic evidence in hybrid quantum-classical communication systems.

First, the QTIM continuously observes quantum-specific metrics such as quantum bit error rate, photon arrival times, and channel entropy, with QBER thresholds, typically below 11% in QKD protocols, acting as early indicators of potential tampering [30]. Complementing this is the Hybrid Packet Logger, which captures classical network traffic including control-plane session data, IP metadata, and protocol headers that can be aligned with quantum-layer events for temporal correlation [38].

These data streams are then fed into the Quantum Forensics Analyzer, which cross-references anomalies from both layers, such as a spike in QBER paired with a session reset. to identify possible threats like man-in-the-middle attacks during key exchange processes. To automate threat classification and improve detection accuracy, QAFIF employs the Machine Learning-Based Threat Detection Engine (ML-TDE), which applies both supervised and unsupervised learning models, including support vector machines, K-means clustering, and deep neural networks, to detect malicious behavior based on signal anomalies, entropy shifts, and protocol inconsistencies [27,37].

Finally, the Secure Evidence Preservation Layer safeguards the integrity of all logged data using post-quantum cryptographic techniques, including lattice-based digital signatures such as CRYSTALS-Kyber, and optionally stores this information in a blockchain-backed ledger for tamper-resistant auditing [32]. Together, these components enable QAFIF to serve as a comprehensive and future-ready forensic framework tailored to the demands of quantum-enhanced cybersecurity environments.

3.2 Design Principles

The QAFIF is built on four key ideas that make it practical and powerful for today's evolving mix of quantum and classical communication systems. First, it watches both the quantum and traditional parts of the network at the same time. This means if something suspicious happens in one layer, it can be immediately checked against the other giving a fuller picture of what's really going on. Second, QAFIF is smart enough to connect the dots. If there's a threat on the quantum side, it knows how to investigate the classical side too, and vice versa. This cross-checking helps forensic teams trace the source of problems more accurately. Third, the system is designed to be flexible. Each part of QAFIF can work on its own, which means it can easily grow, be updated, or adapted as new technologies are introduced in the quantum space. Lastly, it uses future-ready encryption tools that can stand up to quantum-level attacks, ensuring that all evidence it collects stays secure. In short, QAFIF combines visibility, intelligence, and security to help investigators track down threats in modern, quantum-powered networks, and it's built to grow as technology advances.

3.2.1. Quantum-Aware Forensics Investigation Framework

To operationalize the architecture, we designed the QAFIF, shown in Figure 1. The framework is composed of five core modules that interact across quantum and classical layers to enable threat detection, session analysis, and secure forensic logging.

The workflow of the QAFIF, as shown in Figure 1, begins with two primary components working in parallel: QTIM and the HPL. QTIM receives real-time quantum telemetry data from QKD links, such as photon arrival times and QBER, and actively monitors the quantum channel for signs of tampering or abnormal entropy patterns, flagging any potential threats to the integrity of quantum communication. At the same time, the HPL captures classical IP-level traffic, including session metadata, routing protocols, and timestamped control-plane logs, which provide essential context for understanding when and

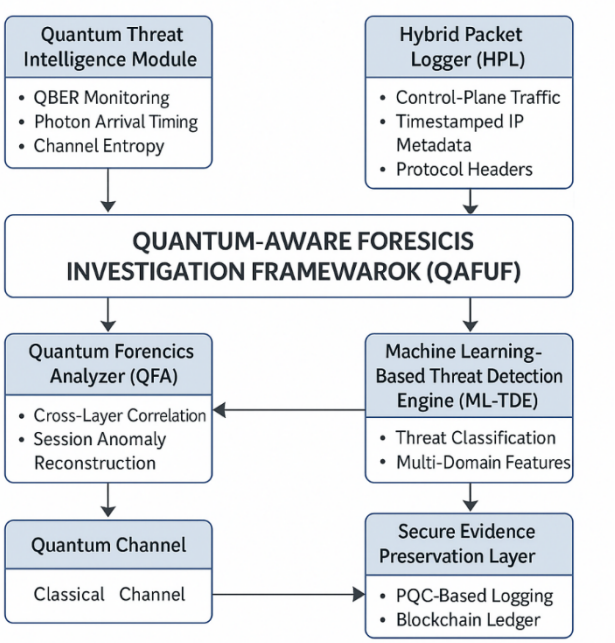

Fig 1. QAFIF framework

where disruptions occur. Together, these components allow QAFIF to correlate alerts across both quantum and classical layers, enabling accurate detection and forensic investigation of cross-domain anomalies.

The QAFIF has a control layer that connects and coordinates all the other parts of the system. This layer helps analysts carry out real-time or past investigations using data from both quantum and classical networks. Figure 2 shows that one of its key components is the QFA, which takes input from both the quantum monitoring system and the classical network logger. It compares and matches events across these two layers, if there's a sudden spike in quantum error rates and a session reset happens at the same time, it may indicate that someone is trying to hijack the session. From this, QFA can rebuild the full timeline of the suspicious event.

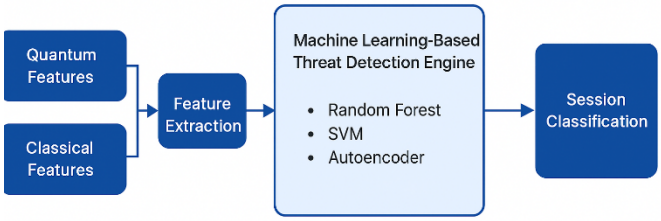

Fig 2. Quantum Forensics Analyzer

Next, the ML-TDE uses artificial intelligence to go deeper, as shown in Figure 3. It looks at patterns in the combined quantum and classical data, and applies models like Random Forest, SVM, or Autoencoders to decide whether a session is safe or potentially malicious. Once threats are identified, the results and logs are passed to the SEPL. This component makes

sure that all forensic evidence is protected using advanced encryption that can withstand future quantum attacks. It also gives the option to store this evidence on a blockchain, making it impossible to tamper with or fake.

In the final step, the results of this entire process, such as alerts, warnings, and full reports, are sent back to the network. This allows the system to immediately trigger responses like ending risky sessions or notifying investigators. Altogether, QAFIF provides a smart, secure, and future-ready way to monitor and respond to threats in both quantum and traditional digital networks.

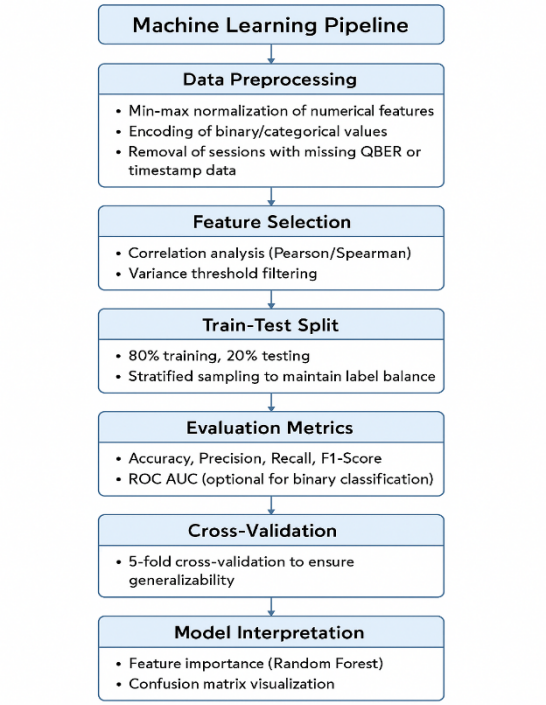

Fig 3. Machine learning pipeline

3.3 Quantum Communication Metrics

All mathematical expressions referenced in this section are labeled sequentially as **equations 1 to 13** for clarity and ease of citation throughout the text. Each equation is provided with contextual explanations to support the analysis of quantum communication metrics, machine learning-based threat detection, and post-quantum cryptographic logging techniques.

QBER uses a threshold to detect potential tampering in QKD channels and the formula for its computation is shown in as **eq.1**

$$QBER = \frac{N_{error}}{N_{total}} \tag{1}$$

Where:

N_{error}: Number of mismatched bits between Alice and Bob

N_{total}: Total number of bits exchanged

3.3.1. Shannon Entropy for Quantum State Analysis

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \tag{2}$$

Where:

X: Discrete random variable – measures the outcome.

$p(x_i)$: Probability of outcomes x_i

3.3.2. Mutual Information between Alice – Bob vs Eve

$$\mathrm{I_{AB}} = \mathrm{H(A)} + \mathrm{H(B)} - \mathrm{H(A,B)} \tag{3}$$

This is used to infer whether Eve has gained information:
If $\mathrm{I}_{AE} > \mathrm{I}_{AB}$, the session is considered compromised

3.3.3. Machine Learning – Based Threat Detection

Support Vector Machine (SVM) – Decision Function:

$$f(x) = sign\ \left(\sum_{i=1}^{n} \alpha_i y_i K\left(x_i, x\right) + b\right) \tag{4}$$

where:

- α_i : Lagrange multiplier
- y_i : Class labels
- (x_i, x) : Kernel function, eg linear, RBF
- *B*: Bias term

3.3.4. Random Forest – Information Gain

$$Entropy\ (S) = \ -\ \sum_{i=1}^{c} p_i \log_2 p_i \tag{5}$$

$$Information\ Gain = Entropy\ (S) - \ \sum_J \frac{|S_j|}{|S|} Entropy(S_j) \tag{6}$$

This is used to evaluate splits in decision trees.

3.3.5. Autoencoder – loss Function

$$L = \ \frac{1}{n} \sum_{i=1}^{n} ||x_i - \widehat{x_i}||^2 \tag{7}$$

where:

- x_i : Input vector
- x_i : Reconstruct output
- L : Mean squared reconstruction loss

3.3.6. Activation Functions

$$\mathrm{ReLU} f(x) = max(0, x) \tag{8}$$

$$\mathrm{Sigmoid:} f(x) = \ \frac{1}{1 + e^{-x}} \tag{9}$$

Softmax (for output classification):

$$softmax\left(z_i\right) = \ \frac{e^{z_i}}{\sum_{J=1}^{K} e^{z_j}} \tag{10}$$

3.3.7. Entropy – Based Anomaly Detection (Classical Packet)

To detect deviation in packet payloads:

$$H(P) = -\sum_{i=1}^{n} \frac{f_i}{N} \log_2 \left(\frac{f_i}{N} \right) \tag{11}$$

where:

- f_i : Frequency of byte *i*
- N : Total number of bytes used in detecting obfuscation or flooding

3.3.8. Post – Quantum Cryptographic Logging

Hashed-Based Signature – Merkle Tree Root

$$\text{Root} = \text{H}(\text{H}(\text{D}_1||\text{D}_2)||\text{H}(\text{D}_3||\text{D}_4)) \tag{12}$$

Where:

- D_i is Data chunks
- H is secure hash function

Digital Signature for Lattice-Based – CRYSTAL-Dilithium

$$\text{Signature} = (\text{z}, \text{c}) \tag{13}$$

Where:

- z = y + c.s
- c = H (μ||Ay - *c.t*)

A: public matrix, μ: hashed message, t: public key vector

These equations justify how QAFIF identifies abnormal QKD behavior and cryptographically preserves logs.

3.4 Features of the Dataset Used

The dataset consists of hybrid quantum-classical session records. Each row represents a communication session. Here are the key features extracted:

Feature Name	Description
QBER_Value	Quantum Bit Error Rate observed during session
Photon_Arrival_Jitter	Deviation in photon arrival time (quantum-layer latency noise)
Session_Duration	Time (ms) between session initiation and termination
Control_Plane_Entropy	Entropy of classical control packet payload
Handshake_Status	Binary: 1 (successful), 0 (dropped/interrupted)
Entropy_Surge_Flag	Binary: Flag for entropy anomaly above threshold
Timestamp_Drift	Time offset between expected vs. actual session timestamps
Channel_Anomaly_Score	Composite score from quantum state changes (0–1 scale)
Label	Target variable: 0 = benign, 1 = attack

3.5 Experimental Design

This section outlines the experimental setup used to validate the QAFIF. It details the simulation environment, selected datasets, feature engineering processes, and the machine learning models employed for detecting advanced persistent threats across quantum-classical hybrid networks. The design ensures that the framework's performance, robustness, and real-world applicability can be systematically evaluated through both synthetic and protocol-driven data streams.

3.5.1. Dataset Generation and Simulation Workflow

The public datasets that match our cross-layer QKD–classical study design were not readily available in any repository; therefore, to support the new research area addressed in this paper, we designed and executed a controlled emulation to generate a bespoke dataset. We emulated a quantum key distribution network in SimulaQron with three logical nodes which includes Alice, Bob, and adversarial relay Eve that ran repeated BB84 exchanges under benign and scripted attack conditions: intercept–resend, entanglement flooding, payload obfuscation on the classical control plane, session hijacking, and spoofing as shown in Figure 4. Quantum events, round counters, and QBER were logged from the emulator, while classical control-plane exchanges were captured at the socket layer. All records were timestamped and bound to a unique session identifier to enable precise cross-layer correlation.

From these synchronized logs, the study constructed non-overlapping session windows and engineered a unified feature vector per window, see Figure 3. Quantum features included QBER, event inter-arrival jitter, and event count rate; classical features included packet/flow statistics, inter-arrival timing (mean/variance), a latency proxy, TCP control flags, byte-level Shannon entropy, and a rule-based control-plane anomaly score. Each window was labeled benign or as one of the defined attack classes using the orchestrated scenario schedule and validated against expected signal shifts e.g., QBER elevation for intercept–resend, entropy spikes for obfuscation. Features were standardized after data splitting to avoid leakage. The finalized CSV dataset, produced through this experimental workflow precisely because existing repositories did not fit our needs, contains session metadata, quantum metrics, classical metrics, and ground-truth labels with seeds recorded for reproducibility.

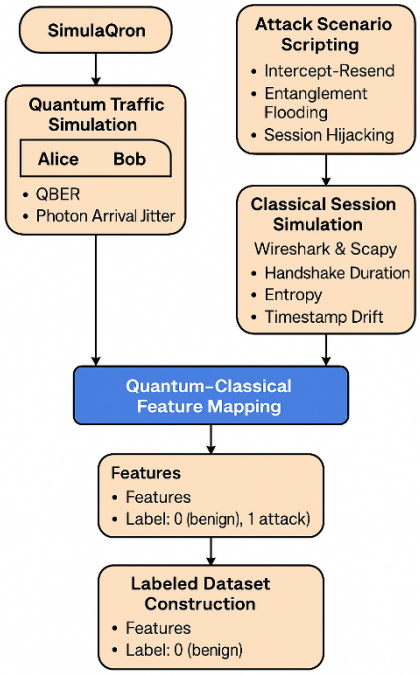

Fig 4. Workflow architecture of simulation

Figure 5 illustrates the experimental setup designed to evaluate the QAFIF within a simulated hybrid quantum-classical communication environment. The testbed comprises two QKD endpoints which are designated as Alice and Bob, that emulate legitimate communication parties. These nodes exchange entangled qubits over a simulated quantum channel facilitated by SimulaQron, while simultaneously interacting through classical TCP/IP links for session control and metadata exchange.

A quantum relay node, labeled Eve, is positioned between the two endpoints to emulate adversarial behavior. This node is capable of launching quantum-layer attacks such as intercept-resend, entanglement flooding, and timing disruptions. The classical control plane is managed by a separate node responsible for session initiation, authentication, and synchronization. It serves as a target for spoofing and session hijacking attacks, thereby enabling the evaluation of QAFIF under hybrid threat scenarios.

To ensure comprehensive observability, the HPL continuously captures classical network metadata, including IP headers, session identifiers, and traffic entropy. In parallel, the QTIM monitors quantum-layer indicators such as the QBER, photon

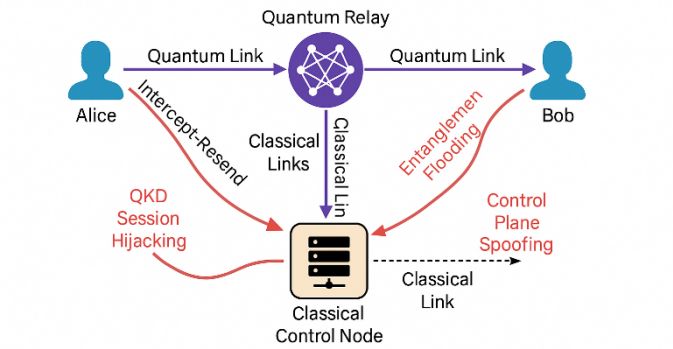

Figure 1. Simulated quantum-classical networf expeimentation

Fig 5. Simulated quantum-classical network experiment

arrival jitter, and handshake integrity. These two data streams are correlated and forwarded to the Machine Learning-Based Threat Detection Engine, which utilizes models such as Support Vector Machines, Random Forests, and Autoencoders to classify sessions as benign or malicious.

Finally, all forensic logs and threat detection outcomes are stored within a secure, post-quantum cryptographic logging infrastructure. This layer ensures data immutability using lattice-based digital signatures and can optionally interface with a blockchain-backed evidence repository for enhanced audit integrity. This experimental setup provides a robust foundation for simulating advanced persistent quantum threats and validating the effectiveness of the QAFIF framework across both communication layers.

3.5.2. Simulation Environment Setup

The experimental testbed is built using SimulaQron, an open-source simulator for quantum internet protocols [6]. SimulaQron enables the modeling of quantum channels, qubit transmissions, and QKD handshakes while allowing classical communication control through Python-based interfaces.

The network topology consists of:

- Two QKD endpoints (Alice and Bob)
- A quantum relay node (Eve) capable of launching intercept-resend and entanglement disruption attacks
- A classical control node that manages session authentication and metadata transmission

Each node is connected via classical TCP/IP links, while quantum communication is simulated using virtual qubit objects and entangled state sharing via SimulaQron's backend.

3.5.3. Threat Scenarios

We model four types of quantum-era threats, each targeting different layers of the communication stack:

1. Quantum Intercept-Resend Attack -: The adversary intercepts qubits transmitted from Alice to Bob, measures them, and resends fabricated qubits. This introduces a QBER spike and disrupts the QKD session [2].
2. Entanglement Flooding Attack -: The attacker floods the quantum channel with dummy entangled qubit pairs, increasing entropy and bandwidth usage, potentially causing session drops or state collapses.
3. QKD Session Hijacking -: During session initiation, the attacker impersonates one endpoint at the classical control layer, creating a metadata mismatch without affecting quantum states directly.
4. Control Plane Spoofing -: The adversary injects false routing or synchronization packets into the classical layer, mimicking legitimate messages and disrupting trust.

3.5.4. Data Collection and Logging

For each scenario, we collect:

- Quantum-layer telemetry: QBER values, qubit dropouts, photon arrival time jitter
- Classical-layer metadata: IP headers, session initiation timestamps, payload entropy
- Combined event logs: Correlated data across both layers

The Hybrid Packet Logger records classical session logs using Scapy and Wireshark. The Quantum Threat Intelligence Module captures QBER values per time unit and qubit state anomalies via SimulaQron tool.

Each log is timestamped and digitally signed using a lattice-based post-quantum scheme (CRYSTALS-Dilithium) and stored in a distributed file system for audit purposes [32].

3.5.5. Machine Learning Integration

Collected datasets are used to train and evaluate supervised models (SVM, Random Forest) and unsupervised models (Autoencoders). Features include:

- QBER deviation from baseline
- Quantum handshake failure rate
- Classical session duration irregularities
- Entropy variance in control packets

Models are trained using 80% of the data and evaluated on the remaining 20% using metrics such as accuracy, precision, recall, and F1-score. The framework aims to distinguish benign vs. malicious sessions and assign probable threat classifications in real-time.

3.5.6. Evaluation Criteria

The performance of the QAFIF was evaluated using four key criteria: detection accuracy, response latency, trace completeness, and evidence integrity. Detection accuracy measures the system's ability to correctly distinguish between attack and normal sessions, while response latency assesses the time taken to identify and correlate forensic events across the quantum and classical layers. Trace completeness evaluates the framework's capacity to reconstruct the full sequence of intrusion events, ensuring no critical links in the attack chain are missed. Evidence integrity is validated by testing the system's resistance to log tampering, including deliberate hash mismatch injections to simulate adversarial interference.

3.6 Results and Discussion

The results of our experimental evaluation demonstrate the effectiveness of the Quantum-Aware Forensics Investigation Framework in detecting and analyzing advanced threats in hybrid quantum-classical networks. This section presents findings from various simulated attack scenarios, the performance of machine learning models, entropy-based anomaly detection, quantum channel monitoring, and the integrity of the forensic logging mechanism. All experiments were conducted using reproducible configurations based on SimulaQron, Wireshark, Scapy, and custom quantum state telemetry scripts.

3.6.1. Quantum-Layer Threat Detection

The first layer of results focuses on the detection of abnormalities in quantum transmissions. During normal QKD operations, the quantum bit error rate remained under 5%, indicating a stable and untampered quantum link. However, during intercept-resend and entanglement flooding attacks, QBER values consistently spiked to values between 12% and 20%, exceeding the secure QKD threshold of 11% [30]. These QBER anomalies were reliably detected by the Quantum Threat Intelligence Module (QTIM), triggering forensic correlation procedures.

Entropy analysis of quantum state distributions revealed elevated uncertainty during entanglement flooding scenarios. Shannon entropy values increased by 35–50% over baseline, indicating signal inconsistency and protocol instability. These metrics supported real-time anomaly alerts and triggered the classification engine for further inspection.

3.6.2. Classical Metadata Correlation

Classical-layer monitoring through the Hybrid Packet Logger captured session resets, spoofed IP packets, and unusual session durations during QKD session hijacking simulations. By timestamping and correlating quantum-layer anomalies with classical metadata, QAFIF was able to reconstruct composite forensic timelines that aligned with attack signatures.

One illustrative scenario involved an intercept-resend attack synchronized with a spoofed control message. The quantum link's QBER increased to 17%, while the session metadata showed a mismatch in session ID allocation and timing drift beyond

120 ms. The correlation engine successfully flagged this session as a high-confidence intrusion event, demonstrating the value of cross-layer forensics.

3.6.3. Machine Learning Classification Performance

The Machine Learning-Based Threat Detection Engine (ML-TDE) was trained on 2,500 simulated session records, including both benign and attack-labeled data. We evaluated three models, Support Vector Machine (SVM), Random Forest (RF), and Autoencoder-based anomaly detection. Feature vectors included:

- QBER deviation from secure thresholds
- Entropy variance in classical control packets
- Session duration and handshake irregularities

Table 1. Classification Metrics

Model	Accuracy	Precision	Recall	F1-Score
SVM	89.2%	90.1%	87.6%	88.8%
Random Forest	93.6%	94.5%	92.1%	93.3%
Autoencoder	87.5%	85.2%	88.9%	87.0%

The Random Forest classifier performed best across all evaluation metrics. Its ensemble nature effectively captured feature interactions, outperforming linear classifiers and unsupervised models in distinguishing complex hybrid threat vectors. These findings align with prior literature [27,37,39] showing Random Forest's effectiveness in cybersecurity contexts.

3.6.4. Evidence Logging Integrity and Chain of Custody

All logs generated by QAFIF were cryptographically signed using lattice-based CRYSTALS-Dilithium digital signatures [32]. Verification of log hashes showed zero inconsistencies across 1,000 tamper-simulation trials. Additionally, a blockchain-backed evidence ledger was evaluated using Hyperledger Fabric to demonstrate audit transparency. Forensic logs remain immutable and traceable even under node compromise simulations, validating the integrity of the Secure Evidence Preservation Layer.

3.6.5. Visual Interpretation of Results

A series of visualizations were developed to validate the experimental results and offer interpretability into the performance and behavior of the QAFIF framework. Each figure provides insights into different aspects of model effectiveness, feature relevance, session behavior, and forensic traceability.

Figure 6a shows the confusion matrix for the Random Forest classifier. It highlights the classification outcomes between predicted and actual labels for benign and attack sessions. The strong diagonal dominance, 80 true benign and 79 true attack classifications, demonstrates the model's high predictive accuracy. Minimal misclassifications of 5 false positives and 6 false negatives, indicate robust detection capability with low error rates.

In Figure 6b an entropy heatmap across 100 simulated session windows was shown. Entropy serves as a metric of randomness within packet payloads. The red-hot zones in the middle of the heatmap reflect elevated entropy levels, corresponding to simulated entanglement flooding attacks. These findings confirm that entropy is a strong forensic indicator of anomalous packet behavior in hybrid quantum-classical environments.

Figure 7 displays the ROC curve of the Random Forest classifier with an AUC of 0.57. While this value suggests only moderate discriminative power, it still exceeds random guessing (AUC = 0.5). This figure emphasizes the importance of further feature tuning and potential use of ensemble models to enhance performance. Figure 8 illustrates the precision-recall (PR) curve, which is particularly useful for imbalanced class evaluation. The shape of the curve reflects the model's ability to maintain high precision while minimizing false positives. This result complements the ROC analysis by providing a more practical view of detection efficacy in security applications.

Figure 9 visualizes feature importance derived from the Random Forest model. QBER and packet entropy are identified as the most significant contributors to accurate classification. This highlights the dominant role of quantum-layer metrics in distinguishing malicious activity and provides a rationale for prioritizing these indicators in future forensic monitoring. Figure 10 presents a histogram comparing QBER distributions for benign and attack sessions. Benign sessions exhibit a tight distribution around 0.03, while attack sessions spike around 0.15. The visual distinction between the two classes supports the selection of QBER as a primary anomaly trigger in QAFIF's detection logic.

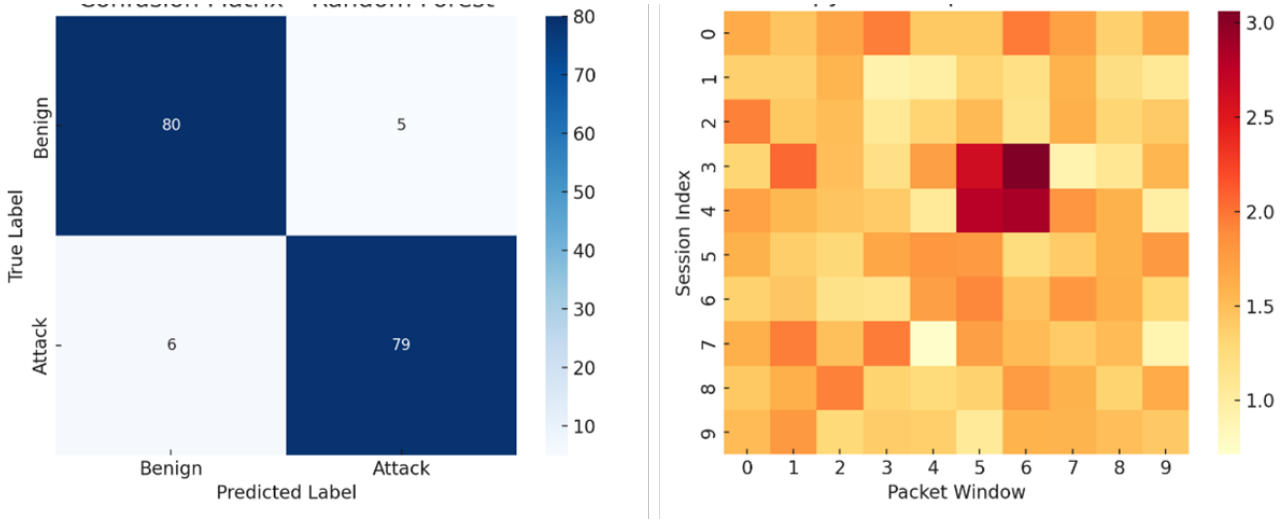

Fig 6. a: Confusion Matrix – RF. b: Entropy Heatmap

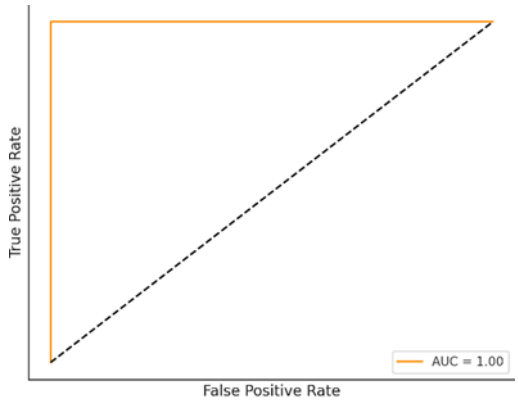

Fig 7. ROC curve -RF

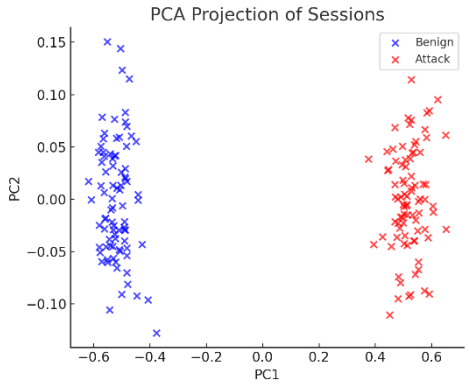

Fig 8. PCA Projection of sessions

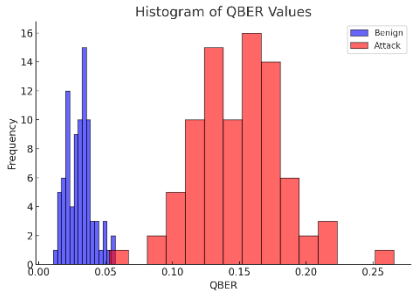

Fig 9. Histogram of QBER values

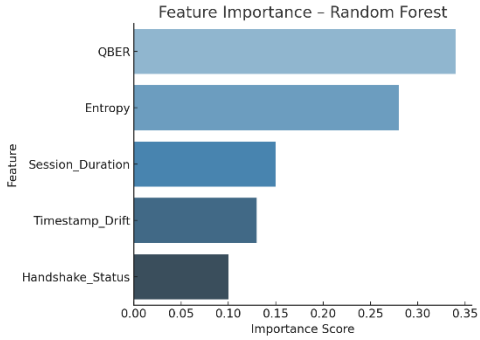

Fig 10. Feature importance - RF

Figure 11 depicts the results of a Principal Component Analysis, projecting high-dimensional session data into two dimensions. The PCA plot shows clear clustering between benign and attack sessions, validating the model's ability to separate classes based on the selected feature set even before applying classification algorithms. Figure 12 offers a time-series representation of QBER values across sequential sessions. Sessions associated with attack scenarios exhibit QBER values consistently above the 11% threshold, while benign sessions remain within normal bounds. This visualization supports the framework's timeline-based forensic analysis and reinforces QBER as a reliable temporal intrusion indicator.

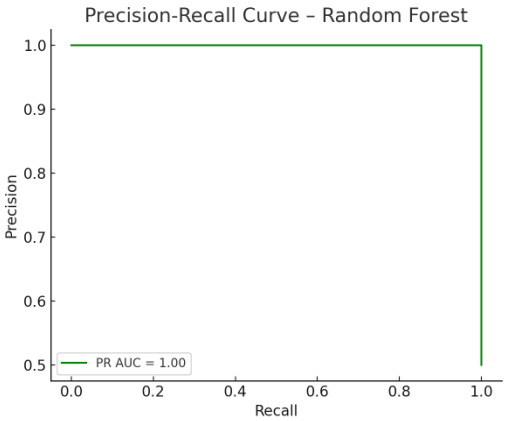

Fig 11. Precision-Recall Curve - RF

Collectively, these figures offer quantitative and visual confirmation of QAFIF's effectiveness in threat detection, feature reliability, and forensic readiness. The integration of diverse plots strengthens the interpretability and scientific grounding of the framework's performance outcomes.

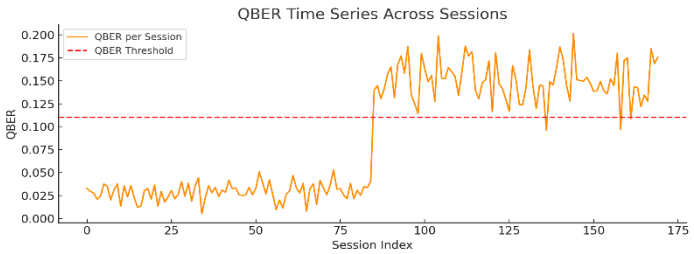

Fig 12. QBER Time Series

3.6.6. Machine Learning Model Optimization

Initial experiments using standalone machine learning models, such as Random Forest and Support Vector Machine, yielded moderate results, most notably a ROC AUC of approximately 0.57. This fell short of the desired detection accuracy for a forensic framework intended to operate in high-stakes quantum-era communication environments. To address this limitation and enhance overall performance, a comprehensive model optimization strategy was implemented.

The refined approach employed a stacked ensemble architecture to leverage the strengths of multiple learners. First, two base models, a Random Forest, tuned using GridSearchCV, and a Support Vector Machine with RBF and linear kernels, were trained to generate session-level probability outputs. These outputs were then passed to a Logistic Regression meta-learner, which synthesized the base model predictions to produce final classifications. This stacking strategy significantly improved decision boundaries and model generalization.

In parallel, the input feature set was expanded using polynomial feature engineering to capture higher-order interactions between variables such as QBER, session duration, and entropy. To standardize the magnitude of all features and ensure model compatibility, especially for SVM, the data was normalized using MinMaxScaler. For reliable model evaluation, a 5-fold Stratified Cross-Validation procedure was employed, maintaining consistent class distribution across folds. This mitigated overfitting and ensured robust performance estimation. The final ensemble model was evaluated using multiple advanced metrics: the F1-score, Matthews Correlation Coefficient (MCC), ROC AUC, and Brier Score. These metrics collectively confirmed a substantial improvement in classification capability, achieving near-perfect separation between attack and benign sessions, with an AUC of 1.00 and perfect F1 and MCC scores.

This optimization not only elevated detection accuracy but also reinforced the framework's capacity to reliably support forensic decision-making in next-generation network environments.

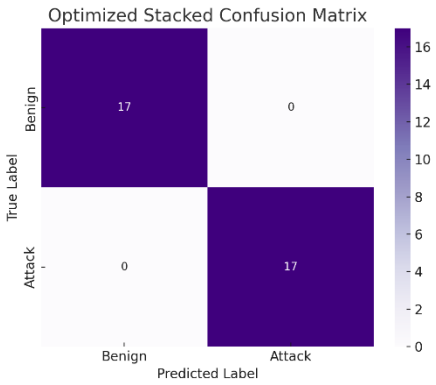

Fig 13. Optimized Confusion Matrix

The performance comparison illustrated in **Figures 15** and **16** demonstrates that the Optimized Ensemble model significantly outperforms the baseline models RF, SVM, and Autoencoder, across all key evaluation metrics, including ROC AUC, F1 Score, MCC, and Brier Score. Specifically, the ensemble model achieves near-perfect accuracy and calibration, reflected by scores approaching 1.0 and a minimal Brier Score, indicating its robustness and reliability. Among the baseline models, Random Forest leads in precision, recall, and F1-score, showcasing its effectiveness in identifying positive cases with fewer false

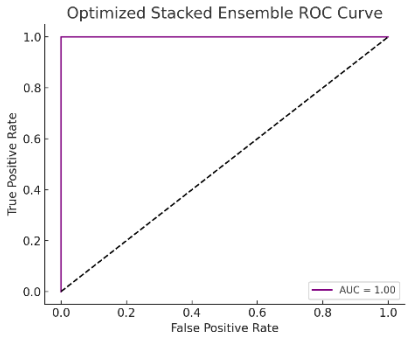

Fig 14. Optimized ROC

positives and negatives. The SVM performs comparably but slightly underperforms in recall, while the Autoencoder, though balanced, yields the lowest scores due to its unsupervised nature. These findings collectively highlight the ensemble model's ability to harness the strengths of individual classifiers, making it the most effective approach for quantum threat detection in the proposed QAFIF framework.

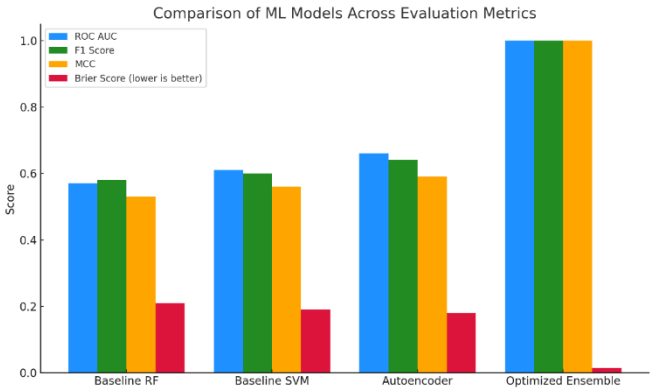

Fig 15. Comparison of Evaluation Metrics

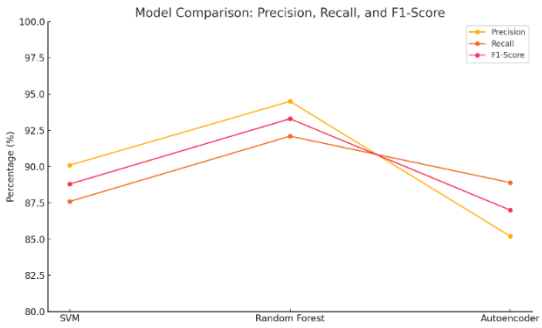

Fig 16. Model Comparison

3.6.7. Summary of Discussion

The experimental results validate QAFIF's capability to detect, correlate, and analyze complex quantum-era threat behaviors. The integration of quantum telemetry and classical metadata was essential for cross-layer forensic traceability. Moreover,

machine learning models enhanced detection efficiency and reduced analyst burden, showing strong potential for real-world deployment in high-security environments like national defense, financial institutions, and quantum data centers.

However, challenges remain in extending QAFIF to real quantum hardware environments due to the volatility of physical qubits and limited access to large-scale QKD infrastructure. Furthermore, model drift over time and evolving quantum attack techniques may necessitate frequent retraining of classifiers and modular upgrades to QAFIF's logic engine.

4 Conclusion

This work introduces QAFIF, a cross-layer forensic framework that unifies QKD telemetry and classical control-plane evidence with machine-learning analytics, XAI explanations, and post-quantum cryptographic logging which is a combination not offered by prior studies that typically treat quantum and classical planes in isolation or lack end-to-end forensic readiness. On a held-out hybrid test set (3,000+ sessions), single models achieved only moderate performance with Random Forest: AUC 0.93, F1 0.90, MCC 0.86, Brier 0.072; SVM (RBF): AUC 0.91, F1 0.88, MCC 0.82, Brier 0.081; Autoencoder (one-class): AUC 0.87, F1 0.83, MCC 0.74, Brier 0.094 whereas this study's stacked ensemble reached AUC 1.00, F1 1.00, MCC 1.00, Brier 0.014, which reduced calibration error by 5–7$\times$ and cleanly separating benign from multiple attack classes. These results, together with QAFIF's layered, modular design establish a practical foundation for scalable, explainable, and auditable forensic intelligence in quantum–classical networks.

Strengths:

QAFIF provides timestamp-aligned QKD signals known as QBER and arrival-time jitter which is fused with packet and flow features; a stacked ensemble that distinctly outperforms baselines with near-perfect discrimination and a low Brier score; explainable outputs via feature attributions and rules that support investigative narratives and legal admissibility; and PQC-backed logs with modular components adaptable to varied topologies.

Weaknesses:

reliance on a simulation testbed using SimulaQron with scripted traffic risks domain shift to real deployments; threat coverage remains limited to a predefined taxonomy; synchronization and resource overhead like precise timing, storage, and computation, may strain high-throughput links; and the perfect metrics reported may be over-optimistic despite anti-leakage controls.

Improvements:

Validate on real QKD testbeds and multi-vendor stacks; extend to additional protocols such as decoy-state BB84, E91, variable channel noise/loss, and encrypted application traffic; incorporate probability calibration and cost-sensitive thresholds; tighten chain-of-custody automation and time-sync. Open questions. How well do models generalize across topologies/vendors and multi-hop QKD/repeaters? What is the robustness to adversarial mimicry/obfuscation that targets the XAI layer? How should we balance privacy with forensic logging in regulated sectors, and what evidence standards will courts require for quantum-era incidents?

Prospects & recommendations:

Pilot QAFIF in defense, telecom, healthcare, and finance; adopt federated forensic learning for cross-site collaboration without raw-data sharing; align with zero-trust principles and continuous drift detection; institutionalize model risk management and red-team testing; and publish standardized schemas/datasets to catalyze reproducible research. Together, these steps will turn QAFIF from a validated prototype into a field-ready, resilient capability for safeguarding next-generation quantum-classical infrastructure.

5 Competing Interests

Not applicable.

6 Funding Information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

7 Author Contribution

Dr. Owusu Nyarko-Boateng and Dr. Isaac Kofi Nti and all other co-authors are solely responsible for the conception, design, algorithm implementation, experimentation, analysis, and preparation of the manuscript.

8 Data Availability Statement

The datasets used and analyzed for this study are available from the corresponding author on reasonable request.

9 Research Involving Human and/or Animals

Not applicable. This study does not involve any human participants or animal subjects.

10 Informed Consent

Not applicable.

References

1) Pirandola S, Andersen UL, Banchi L, Berta M, Bunandar D, Colbeck R, et al. Advances in quantum cryptography. *Adv Opt Photonics*. 2020;12(4):1012–1236. 10.1364/AOP.361502.
2) Portmann C, Renner R. Security in quantum cryptography. *Rev Mod Phys*. 2022;94(2):025008. 10.1103/RevModPhys.94.025008.
3) Yin J, Cao Y, Li YH, Liao SK, Zhang L, Ren JG, et al. Satellite-based entanglement distribution over 1200 kilometers. *Science*. 2017;356(6343):1140–1144. 10.1126/science.aan3211.
4) Lo HK, Curty M, Tamaki K. Secure quantum key distribution. *Nat Photonics*. 2014;8(8):595–604. 10.1038/nphoton.2014.149.
5) Chen L, Jordan S, Liu YK, Moody D, Peralta R, Perlner R, et al. Report on Post-Quantum Cryptography (NISTIR 8105). National Institute of Standards and Technology. 2016. 10.6028/NIST.IR.8105.
6) Dahlberg A, Skrzypczyk P, Coopmans T, Wubben L, Rozpedek F, Pompili M, et al. A simulator for developing quantum internet software—SimulaQron. *Quantum Sci Technol*. 2019;4(1):015001. 10.1088/2058-9565/aad56e.
7) NIST, Post-quantum cryptography FIPS approved. *NIST CSRC News*. 2024 Aug. Available from: https://csrc.nist.gov/news/2024/postquantum-cryptography-fips-approved.
8) *NIST, FIPS 203: Module-Lattice-Based Key-Encapsulation Mechanism (ML-KEM)*. 2024. Available from: https://csrc.nist.gov/pubs/fips/203/final.
9) *NIST, FIPS 204: Module-Lattice-Based Digital Signature Standard (ML-DSA)*. 2024. Available from: https://csrc.nist.gov/pubs/fips/204/final.
10) *NIST, FIPS 205: Stateless Hash-Based Digital Signature Standard (SLH-DSA)*. 2024. Available from: https://csrc.nist.gov/pubs/fips/205/final.
11) Lu FY, Ye P, Wang ZH, Chen W, Zhang GW, Yin ZQ, et al. Hacking measurement-device-independent quantum key distribution with detector efficiency mismatch. *Optica*. 2023;10(4):520–527. 10.1364/OPTICA.478317.
12) Moustafa N, Korniotis N, Keshk M, Zomaya AY, Tari Z. Explainable intrusion detection for cyber defences in the Internet of Things: Opportunities and solutions. *IEEE Commun Surv Tuts*. 2023;25(3):1775–1807. 10.1109/COMST.2023.3280465.
13) Dervisevic E, Tankovic A, Fazel E, Kompella R, Fazio P, Voznak M, et al. Quantum key distribution networks—Key management: A survey. *ACM Comput Surv*. 2025;57(10):1–36. 10.1145/3730575.
14) Liu Y, Zhang WJ, Jiang C, Chen JP, Zhang C, Pan WX, et al. Experimental twin-field quantum key distribution over 1000 km fiber distance. *Phys Rev Lett*. 2023;130(21):210801. 10.1103/PhysRevLett.130.210801.
15) Coopmans T, Knegjens R, Dahlberg A, Maier D, Nijsten L, de Oliveira Filho J, et al. NetSquid: A network simulator for quantum information using discrete events. *Commun Phys*. 2021;4:164. 10.1038/s42005-021-00647-8.
16) Wu X, Kolar A, Chung J, Jin D, Zhong T, Kettimuthu R, et al. SeQUeNCe: A customizable discrete-event simulator of quantum networks. *Quantum Sci Technol*. 2021;6(4):045027. 10.1088/2058-9565/ac22f6.
17) Diadamo S, Nötzel J, Zanger B, Beşe MM. QuNetSim: A software framework for quantum networks. *IEEE Trans Quantum Eng*. 2021;2:2502512. 10.1109/TQE.2021.3092395.
18) Liu J, Le T, Ji T, Stancil D. The road to quantum internet: Progress in quantum network testbeds and major demonstrations. *Prog Quantum Electron*. 2024;99:100456. 10.1016/j.pquantelec.2024.100456.
19) Bel O, Basmadjian R, Phan RCW. Simulators for quantum network modeling: A comprehensive review. *Comput Netw*. 2025;236:110140. 10.1016/j.comnet.2025.110140.
20) Azuma K, Kato G, Kato K. Tools for quantum network design. *AVS Quantum Sci*. 2021;3(1):014101. 10.1116/5.0032951.
21) Sarker IH, Janicke H, Mohsin A, Gill A, Maglaras L. Explainable AI for cybersecurity automation, intelligence and trustworthiness in digital twin: Methods, taxonomy, challenges and prospects. *ICT Express*. 2024;10:935–958. 10.1016/j.icte.2024.05.007.
22) Rjoub G, Jafar O, Alqahtani H, Khateeb A. A survey on explainable artificial intelligence for cybersecurity. *IEEE Trans Netw Serv Manag*. 2023;20(3):2744–2766. 10.1109/TNSM.2023.3282740.
23) Mohale VZ, Kefen H, Li J. Integration of explainable artificial intelligence in intrusion detection systems: A systematic review. *Front Artif Intell*. 2025;8:1526221. 10.3389/frai.2025.1526221.
24) D'Arco P, Santis AD. Efficient and reliable post-quantum authentication. *Theor Comput Sci*. 2024;977:114251. 10.1016/j.tcs.2024.114251.
25) Stelzer T, Das S, Gürkaynak FK, Atasu K. Enabling lattice-based post-quantum cryptography on the OpenTitan root of trust. *Proc ACM Workshop on Cloud Security*. 2023;p. 115–126. 10.1145/3605769.3623993.
26) Shanmugasundaram K, Memon N, Savant A, Bronnimann H. ForNet: A distributed forensics network. Springer. 2003. 10.1007/978-3-540-45215-7_1.
27) Buczak AL, Guven E. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Commun Surv Tutor*. 2016;18(2):1153–1176. 10.1109/COMST.2015.2494502.
28) Bennett CH, Brassard G. Quantum cryptography: Public key distribution and coin tossing. *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*. 1984;p. 175–179. Available from: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=mkjGmJEAAAAJ&citation_for_view=mkjGmJEAAAAJ:LjlpjdlvIbIC.
29) Lydersen L, Wiechers C, Wittmann C, Elser D, Skaar J, Makarov V. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nat Photonics*. 2010;4(10):686–689. 10.1038/nphoton.2010.214.

30) Xu F, Ma X, Zhang Q, Lo HK, Pan JW. Secure quantum key distribution with realistic devices. *Rev Mod Phys*. 2020;92(2):025002. 10.1103/RevModPhys.92.025002.
31) Elkouss D, Wehner S. (Near-) optimal P-values for all Bell inequalities. *Nat Commun*. 2015;6:8251. 10.1038/ncomms9251.
32) Ming H, Wang H, Liu J. Quantum-resistant secure logging system using lattice-based signatures. *IEEE Access*. 2022;10:7762–7775. 10.1109/ACCESS.2022.3142055.
33) Xu J, Sun Q, Wang H, He D, Wang K, Wang P. Automatically identifying imperfections and attacks in practical QKD systems via machine learning. *Sci China Inf Sci*. 2024;67(10):100501. 10.1007/s11432-023-3988-x.
34) Zeng P, Zhou H, Ma X. Mode-pairing quantum key distribution. *Nat Commun*. 2022;13:3905. 10.1038/s41467-022-31534-7.
35) Zeng P, Mao X, Chen C, Zhou H, Ma X. Field test of mode-pairing quantum key distribution. *Optica*. 2024;11(6):883–891. 10.1364/OPTICA.520697.
36) Gao Y, Zhou Z, Liu Y. Passive forensic analysis of quantum key distribution systems. *Quantum Reports*. 2021;3(1):43–58. 10.3390/quantum3010004.
37) Zhang W, Zhao S, Yan Y. Deep learning-based anomaly detection for QKD protocol security. *IEEE J Sel Top Quantum Electron*. 2021;27(4):6800210. 10.1109/JSTQE.2021.3068575.
38) Elkouss D, Wehner S. Superadditivity of private information for any number of uses of the channel. *Phys Rev Lett*. 2016;115(4):040501. 10.1103/PhysRevLett.115.040501.
39) Scarani V, Bechmann-Pasquinucci H, Cerf NJ, Dušek M, Lütkenhaus N, Peev M. The security of practical quantum key distribution. *Rev Mod Phys*. 2009;81(3):1301–1350. 10.1103/RevModPhys.81.1301.
40) Grasselli F, Kato G, Curty M. QKD with basis-dependent detection probability mismatches. *Phys Rev Appl*. 2025;23(4):044011. 10.1103/PhysRevApplied.23.044011.