RESEARCH ARTICLE

DOI: 10.54392/irjmt26326

INTERNATIONAL RESEARCH JOURNAL OF MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

Quantum Key Distribution (QKD) Enabled Secure Model Update Exchange in Federated Learning

Venkadeshan Ramalingam [a], Basant Kumar [b], S. Jagadeesan [c,*], T. Vetriselvi [d], T. Sivakumar [e]

[a] Lincoln University College, Petaling Jaya, Selangor Darul Ehsan, Malaysia
[b] Modern College of Business & Science, Oman
[c] Department of Software and Systems Engineering, Vellore Institute of Technology, India
[d] Department of Internet of Things, School of Computer Science and Engineering, Vellore Institute of Technology, India
[e] Department of Analytics School of Computer Science and Engineering, Vellore Institute of Technology, India
* Corresponding Author Email: s.jagadeesan@vit.ac.in

DOI: https://doi.org/10.54392/irjmt26326
Received: 08-12-2025; Revised: 22-04-2026; Accepted: 15-05-2026; Published: 30-05-2026

Abstract: Federated Learning (FL) enables collaborative model training while preserving data locality, but the exchange of model updates between distributed clients and a central server remains vulnerable to interception and manipulation. This paper introduces a secure federated learning communication framework that integrates Quantum Key Distribution (QKD) using the BB84 protocol for secure key establishment. The proposed framework uses QKD-generated symmetric keys together with classical cryptographic mechanisms to protect the confidentiality and integrity of model-update messages. A system-level simulation environment is developed to model communication, key generation, authentication, and encryption overheads in federated learning. The simulation considers model-update size, number of clients, communication latency, and key-refresh cost to evaluate system performance and feasibility. The results show that introducing QKD adds limited and predictable overhead to communication latency and key management while preserving model convergence. This study focuses on communication-layer security and does not address learning-layer threats such as adversarial model poisoning or malicious client behavior. Overall, the findings indicate that quantum-secured key exchange can be integrated into federated learning systems and may serve as an enabling technology for secure communication in future distributed learning environments.

Keywords: Quantum Key Distribution (QKD), Federated Learning (FL), Quantum Cryptography, Distributed Machine Learning, Privacy Preservation, Secure Aggregation, Quantum-Resilient Communication.

1. Introduction

Federated Learning (FL) has rapidly matured into a practical paradigm for distributed model training that preserves data locality and user privacy by keeping raw data on edge devices and only sharing model updates with a central aggregator [1]. FL has been adopted through mobile applications, IoT/edge systems, and healthcare use-cases where regulatory or privacy constraints prohibit centralizing sensitive data. Recent surveys summarize FL's advantages and the broad set of application domains where it is being deployed [2]. Despite its privacy-preserving design, FL is vulnerable to a range of security and privacy threats that arise from the transmission and aggregation of local model updates. Attacks such as model/gradient inversion, membership inference, poisoning (including backdoor) and Byzantine-style manipulations can leak private information or corrupt the global model [3].

Further, the communication channel used to transfer model updates makes federated learning systems vulnerable to interception, replay, and man-in-the-middle attacks unless adequate cryptographic protection is applied. Several studies [4] and proposals, including secure aggregation and post-quantum secure aggregation schemes, document these vulnerabilities and possible mitigation strategies. At the same time, rapid advances in quantum computing pose a serious threat to many classical public-key primitives, such as RSA and ECC, that currently underpin secure communications. This creates the "harvest-now, decrypt-later" problem for long-lived or high-value data, in which adversaries may store encrypted model updates today and decrypt them once sufficiently powerful quantum computers become available [5]. A multi-layered post-quantum secure machine-learning framework based on lattice cryptography has been proposed to protect distributed AI systems [6],

emphasizing end-to-end security across training and inference stages in edge- and cloud-based ML environments. Another path toward quantum-resilient security is Quantum Key Distribution (QKD), which relies on quantum-physical principles rather than computational hardness to generate symmetric keys with information-theoretic security. Prior work has introduced dynamic on-demand key allocation for QKD-based IoT networks to improve secure communication in quantum internet environments while maintaining efficient key management [7]. By combining QKD with one-time-pad or authenticated symmetric encryption, retrospective decryption can be made resistant even to adversaries with very high computational power. These developments motivate the central hypothesis of this work: integrating QKD into the FL communication channel for model-update exchange can significantly improve the confidentiality and integrity of updates, protect FL against current network attacks and future quantum-enabled threats, and remain feasible for smart-environment systems [8]. Related work has also presented architectures that combine distributed sensing systems with quantum communication networks to support secure and coordinated data exchange [9]. However, QKD typically depends on dedicated quantum channels, such as optical fiber, free-space optical links, or trusted-node relays, which may be impractical in heterogeneous IoT settings and introduce significant engineering trade-offs. In addition, a post-quantum blockchain-based federated learning protocol (PQBFL) has been proposed to improve privacy, trust, and resistance to quantum-enabled threats in distributed AI training [10]. Because federated learning often requires frequent transmission of large model updates, the QKD key-generation rate and associated classical post-processing must be sufficient to protect update traffic without introducing prohibitive latency [11]. Protocol design choices such as key buffering, hybrid or pure symmetric keys, and constrained key-reuse schemes therefore require careful balancing of security and performance. A comprehensive threat model must also consider classical network adversaries, quantum adversaries, and compromised or Byzantine clients. Accordingly, QKD should be combined with complementary defenses, such as secure aggregation, anomaly detection, and robust aggregation rules, to cover the broader attack surface [12].

In this paper we propose and analyze a QKD-Enabled Secure Model Update Exchange framework for FL. Our main contributions are: (i) we design a practical architecture that integrates QKD key establishment between clients (or between clients and aggregator nodes) and the FL control/aggregation plane. The architecture supports hybrid key use (symmetric encryption + message authentication) to protect updates while respecting key-rate limits. (ii) We present protocols for key negotiation, update encryption/authentication, and intrusion detection (leveraging QKD's eavesdropping detection signals) mapped to FL training rounds. (iii) We analyze how QKD increases resilience against eavesdropping and retrospective decryption and how it composes with defenses against poisoning and inference attacks. The remainder of the paper is organized as follows. Section 2 provides a comprehensive literature survey on federated learning security, post-quantum defenses, and developments in quantum key distribution. Section 3 presents the proposed methodologies, including the system design, threat model, and QKD-enabled protocols for secure model update exchange. Section 4 details the experimental setup, simulation parameters, and discusses the results obtained from performance and security evaluations. Finally, Section 5 concludes the paper and outlines directions for future research.

2. Related Works

The security and privacy weaknesses of traditional Federated Learning (FL) have been widely studied. A recent survey categorizes threats into model and data attacks (e.g., poisoning, backdoor, and Byzantine attacks), privacy attacks (e.g., inference and reconstruction), and communication-related vulnerabilities caused by interception or tampering of model updates [13]. Because FL often operates over untrusted networks and involves many distributed participants, such as clients, edge devices, and IoT nodes, transmitted model updates, gradients, or parameter deltas are susceptible to man-in-the-middle, replay, and eavesdropping attacks, as well as malicious updates from compromised clients [10]. Moreover, even when cryptographic protections such as secure channels, classical encryption, and digital signatures are used, the rise of quantum computing challenges their long-term security [14]. Several recent studies argue that traditional public-key cryptosystems may become insecure under future quantum attacks, creating the so-called "harvest-now, decrypt-later" threat [15]. Thus, although FL offers important benefits in data locality and privacy preservation, securing model-update communication remains a critical and nontrivial problem, particularly in the presence of future quantum threats.

To address quantum threats, recent research has explored the integration of post-quantum cryptography into FL protocols. One example is PQSF: Post-Quantum Secure Privacy-Preserving Federated Learning, which uses secret sharing and double masking to protect model parameters and updates, thereby enabling quantum-resistant secure aggregation [16]. More recently, blockchain-assisted federated learning architectures based on post-quantum cryptographic signatures have been proposed to strengthen identity protection and model-exchange security against future quantum attacks [17].

Table 1. Summary of Related Works

Paper ID	Approach	Quantum-Safety	Pros	Cons / Gaps
[1]	Comprehensive survey on FL privacy, poisoning, inference, and communication attacks	Not quantum-safe	Strong foundation for FL security research; identifies key vulnerabilities	Does not address quantum threats or QKD/PQC integration
[2]	Cryptographic secure aggregation protocol for federated learning	Not quantum-safe	Foundational secure aggregation model for FL	Classical cryptography vulnerable to quantum attacks
[3]	Survey of defense mechanisms for FL	Not quantum-safe	Covers poisoning, privacy leakage, and adversarial defense	Limited focus on post-quantum or QKD-based solutions
[4]	Systematic survey of FL architectures and optimization	Not quantum-safe	Broad overview of FL technologies and applications	Security discussion lacks quantum-resilient mechanisms
[5]	Privacy-enhanced FL for IoT infrastructures	Partially quantum-safe	Improves IoT anomaly detection while preserving privacy	Vulnerable to future quantum cryptanalysis
[6]	Layered security architecture integrating post-quantum cryptographic (PQC)	Quantum-safe	Modular multi-layer design improves flexibility and End-to-end ML security coverage	High computational overhead due to PQC
[7]	QKD network with dynamic on-demand key allocation	Quantum-safe	Efficient dynamic key allocation improves resource usage	Scalability challenges in large IoT deployments and Distance limitation of QKD links
[8]	Key-homomorphic pseudo-random function for secure FL	Quantum-resistant	Efficient cross-silo FL communication	Focused only on aggregation layer security
[9]	Integrated architecture combining distributed sensing systems with quantum communication networks	Quantum-safe	Improves coordination in distributed quantum networks	High infrastructure and deployment cost
[10]	Integrate FL, Blockchain, and Post-Quantum cryptography	Quantum-resistant	Strong hybrid security	Communication latency in FL aggregation
[11]	Enhanced FL framework for healthcare collaboration	Partially quantum-safe	Supports privacy-preserving medical AI collaboration	No integration with QKD or quantum-secure channels
[12]	High-dimensional QKD over multicore fiber	Quantum-safe	High key generation rate and long-distance secure communication	Requires specialized optical infrastructure
[13]	Comprehensive FL challenges and research directions	Not quantum-safe	Widely referenced foundational FL study	No quantum-resilient communication perspective
[14]	Post-quantum secure aggregation using lattice-based cryptography	Quantum-resistant	Protects aggregation process against quantum adversaries	Increased computational overhead and key management complexity

[15]	Analysis of FL attack vectors and mitigations	Not quantum-safe	Updated overview of emerging FL threats	Quantum attack resilience not fully addressed
[16]	Privacy-preserving FL with post-quantum security	Quantum-resistant	Combines privacy preservation and PQC robustness	Communication and computation overhead remain high
[17]	Blockchain Federated Learning with PQC signatures (ML-DSA 65) and smart-contract verification	Partially quantum-safe	Quantum resistant authentication, integrity checks, transparency and low cryptographic overhead	The blockchain overhead, size of the signature, and the fact that the consensus layer is not entirely quantum-safe are all significant challenges
[18]	Variational Quantum Circuit (VQC)-enhanced Federated Deep Reinforcement Learning with Post-Quantum Cryptography	Quantum-safe	High security for 6G/IoT environments, strong resistance to quantum attacks, reduced communication latency	High computational overhead due to hybrid quantum-classical simulation
[19]	Review of QKD communication technologies	Quantum-safe	Detailed overview of QKD deployment architectures	Does not integrate QKD with FL environments
[20]	Clustered Federated Learning integrated with hybrid quantum-enhanced optimization and privacy-preserving aggregation	Partially quantum-safe	Strong privacy preservation for sensitive data	Increased communication overhead, limited real-world quantum deployment feasibility
[21]	Device-independent QKD mechanisms	Quantum-safe	Strong security even with untrusted devices	Practical implementation remains difficult and costly
[22]	PQC-based cybersecurity framework with nature-inspired adaptive defense techniques	Partially Quantum-Safe	Enhances quantum-threat readiness, supports cryptographic agility, improves adaptive threat detection and response	Primarily conceptual, limited real-world validation, implementation complexity, and PQC computational overhead
[23]	Personalized Quantum Federated Learning using adaptive client-specific model aggregation	Quantum-resistant	Improves personalization of global models for heterogeneous clients	Limited real quantum hardware deployment feasibility
[24]	Lightweight Post-Quantum Cryptography framework integrating lattice-based and hash-based cryptographic schemes	Quantum-Safe	Strong resistance against quantum computing attacks	Reduced performance under large model sizes
[25]	Quantum-assisted FL for IoT and 6G environments	Partially quantum-safe	Explores future-ready IoT security architecture	Conceptual framework with limited implementation details

Although these frameworks improve trust, traceability, and tamper resistance in federated learning systems, they still face open challenges related to blockchain scalability, transaction cost, and fully quantum-secure consensus. These studies show that PQC techniques, including lattice-based encryption, signatures, secure aggregation, and masking, are viable for enhancing FL security and privacy in a post-quantum environment. However, they remain largely within the classical communication paradigm and do not exploit quantum-physical mechanisms such as QKD or quantum channels; as a result, they do not provide information-theoretic security or quantum-based intrusion detection. Other work has proposed a quantum federated learning framework integrated with post-quantum cryptography for distributed learning in 6G-enabled IoT networks, using lattice-based cryptographic mechanisms such as CRYSTALS-Kyber together with variational quantum circuits to improve learning efficiency and robustness against quantum-enabled adversaries [18]. Overall, the literature suggests growing interest in combining quantum cryptographic protocols, including QKD, quantum homomorphic encryption, and quantum communication, with FL to address quantum threats while leveraging quantum-specific advantages. Nevertheless, this research area remains limited, and most proposals are still at the conceptual or early experimental stage rather than widely deployed [19]. On the quantum-communication front, practical advances in Quantum Key Distribution (QKD) have significantly improved the feasibility and robustness of real-world secure communication. The survey “Application and Development of QKD-Based Quantum Secure Communication” reports that QKD protocols can provide information-theoretic security, interoperate with classical networks, and are gradually becoming suitable for industrial deployment.

Therefore, there remains a significant research gap: a practical, well-specified architecture that uses QKD for secure model update exchange in FL bridging quantum-secure communication and distributed learning has not yet been thoroughly studied. Our proposed work aims to fill this gap, by designing and analyzing such a framework: combining QKD-based key exchange, symmetric encryption for model updates, and secure FL aggregation, thereby delivering quantum-resilient confidentiality, integrity, and quantum-attack detection, while considering practical constraints (key-rate, network topology, scalability). Based on the above analysis, the key research gap lies in the lack of system-level frameworks that integrate QKD into federated learning and rigorously evaluate its communication overhead and feasibility constraints (Table 1).

This work addresses this gap by:

- Proposing a QKD-enabled secure communication architecture for federated learning
- Modeling the interaction between key generation, communication cost, and training rounds
- Providing a simulation-based evaluation of overhead and scalability limits

Unlike prior work [19] that focuses primarily on cryptographic design, this study emphasizes practical system integration and performance trade-offs, thereby offering a complementary perspective on securing federated learning in the presence of quantum-era threats.

3. Methodology

This section outlines the proposed QKD-enabled secure model-update exchange framework for Federated Learning (FL). The approach combines Quantum Key Distribution, post-quantum cryptographic primitives, robust aggregation schemes, and secure exchange of local model updates between distributed clients and the central server to ensure confidentiality, authenticity, and resilience.

3.1 System Model

Figure 1 presents the system model, which consists of the following components:

- ***FL Server (Aggregator):*** Maintains the global model, coordinates training rounds, distributes model parameters, and aggregates authenticated client updates.
- ***FL Clients (Edge Devices):*** Participate in collaborative training by computing local gradients on private datasets and sending encrypted updates to the server.
- ***QKD Infrastructure*:** Enables secure distribution of symmetric keys between clients and the server using quantum channels (BB84 protocols).
- ***Classical Authentication and Communication Channels:*** Used for transmitting model updates, metadata, and aggregated results.

The system follows synchronous FL, where each communication round involves simultaneous client update submissions followed by secure model aggregation [20].

3.1.1 Communication-Layer Guarantees

QKD ensures confidentiality and integrity of model updates during transmission by enabling secure key exchange and detecting interception or man-in-the-middle attempts on the quantum/classical channels.

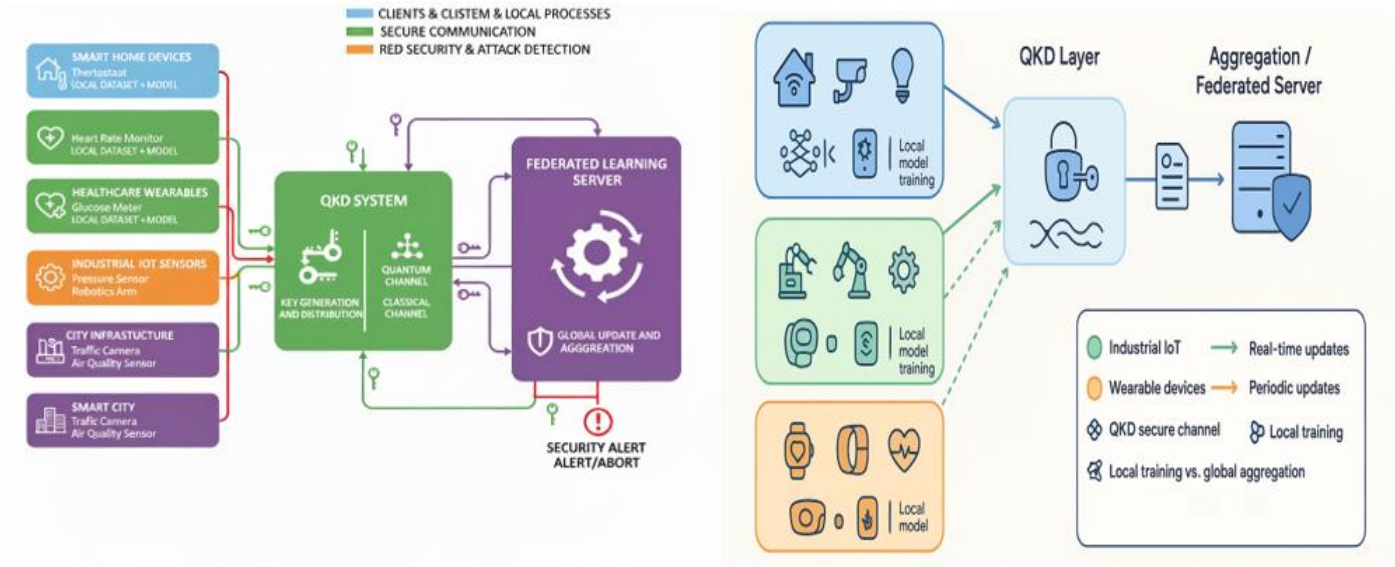

Figure 1. QKD based secure model aggregation

3.2. Federated Learning-Layer Robustness

A privacy-preserving hybrid federated learning framework for mental healthcare has been reported by integrating clustered learning with quantum-enhanced techniques. That approach improves data privacy and security through decentralized model training while using clustering to address data heterogeneity across clients, and it reports improved model accuracy and robustness in sensitive healthcare environments [20]. In the present framework, QKD is used to protect the confidentiality and integrity of transmitted model updates and to provide indirect evidence of eavesdropping through the quantum bit error rate (QBER). As a result, passive interception and some forms of man-in-the-middle attack during key exchange can be addressed at the communication layer. However, QKD does not address adversarial behavior at the learning layer. Threats such as malicious client participation [21], model poisoning, data poisoning, and compromised endpoint devices remain outside the scope of QKD because they target the training process rather than the communication channel. To address this limitation, the proposed framework treats QKD as a secure communication enabler and relies on complementary mechanisms, including robust aggregation, anomaly detection, and trust- or reputation-based client selection, to strengthen learning-layer robustness.

3.3 Quantum Key Distribution Phase

Before each training session, the server performs Quantum Key Distribution (QKD) with the participating clients using the BB84 protocol. In this process, qubits are transmitted over a quantum channel, and the measurement bases are later compared over an authenticated classical channel to generate a shared raw key. The raw key is then processed through error-correction methods such as Cascade to resolve mismatches caused by noise or possible eavesdropping, followed by privacy-amplification methods such as universal hashing to remove partially leaked information. The resulting distilled key can then be used as a symmetric session key. In the proposed framework, AES-256 is used to encrypt model updates, while HMAC with SHA-3 is used to ensure data integrity and authentication. These keys are refreshed for each federated learning round, subject to the available QKD key-generation rate.

3.4 Key Generation Process

Key generation follows the standard end-to-end flow of the BB84 protocol, including quantum-state preparation through polarization encoding, transmission over a quantum channel, and measurement at the receiver using randomly selected bases. The selected bases are then reconciled over an authenticated classical channel to produce a shared raw key [22]. Classical-channel authentication is implemented using pre-shared symmetric keys and message authentication codes (MACs) to mitigate man-in-the-middle attacks. Session keys are refreshed dynamically at the end of each federated learning round based on thresholds such as the quantum bit error rate (QBER) and the average key-usage rate, thereby supporting continuous security and efficient key utilization [23].

3.5 Local Training Phase

In the local training phase, each client privately trains on its own dataset using a global model provided by the server, which includes the initial model parameter initialization; the client then performs training over its local dataset D_i for E epochs, computes the model update Δw_ior the updated model $w_i^{(t+1)}$, and ensures that all local data and intermediate computations remain

securely stored on the device without being shared, thereby preserving data confidentiality throughout the process.

3.6 Secure Model Update Packaging

Before uploading the model update, each client performs the following steps:

The model update is encrypted using the QKD-derived key:

$$C_i = \text{AES-256-CTR}(K_i^{qkd}, \Delta w_i) \quad (1)$$

A message authentication code is computed:

$$MAC_i = \text{HMAC-SHA3}(K_i^{qkd}, C_i \,||\, \text{metadata}) \quad (2)$$

The encrypted package $P_i = (C_i, MAC_i)$ is transmitted to the server over the classical channel.

3.7 Server-Side Verification and Secure Aggregation

To evaluate resilience against adversarial behavior, the proposed framework compares several aggregation strategies, including conventional Federated Averaging (FedAvg) and more robust alternatives. FedAvg computes a weighted average of client updates and is computationally efficient, but it is vulnerable to adversarial manipulation. To reduce this weakness, stronger aggregation techniques are also considered. Krum selects the client update that is closest to its neighbors in Euclidean space, which helps suppress outliers when only a small number of clients are malicious. Trimmed Mean further improves robustness by removing extreme coordinate values before averaging, thereby reducing the impact of adversarial noise. When the server receives client updates, it first verifies integrity and authenticity by recomputing the Message Authentication Code (MAC) using keys derived through Quantum Key Distribution (QKD). Any update that fails verification is rejected to reduce the risk of tampering, poisoning, or man-in-the-middle interference during transmission.

Decryption: Valid updates are decrypted:

$$\Delta w_i = \text{AES-256-CTR}^{-1}(K_i^{qkd}, C_i) \quad (3)$$

Strong Aggregation: In order to counter-poisoning, the server employs a strong aggregator like: Krum, Trimmed Mean, Media aggregation, Norm bounding. The aggregate update is calculated.

$$w^{t+1} = w^t + \eta \cdot \text{Aggregate}(\{\Delta w_i\}) \quad (4)$$

3.8 Global Model Update Distribution

At the end of each federated learning round, the updated global model is securely shared with all participating clients over encrypted classical communication channels. Transmission can be protected using keys generated through Quantum Key Distribution (QKD) or, where appropriate, post-quantum public-key encryption (PQ-PKE) schemes to ensure confidentiality and resistance to both classical and quantum adversaries. This secure distribution preserves the integrity of the transmitted model parameters and allows clients to synchronize with the latest global model state. Once the updated model is delivered, clients can begin the next stage of local training, thereby completing one full federated learning cycle.

3.9 Overhead and Feasibility Model for QKD-Enabled Federated Learning

To examine the scalability of the proposed framework, we model the relationship between federated learning communication requirements and the key-generation capability of the QKD system.

Algorithm: QKD-Enabled Secure FL Training

Input: Global model w^0, client datasets D_i, rounds T

Output: Final global model w^T

1. **QKD Phase:**
2. For each client i, perform QKD → obtain key K_i^{qkd}
3. **For each round t from 0 to T–1 do:**

Client Side:

a. Receive global model w^t

b. Train locally → compute update Δw_i^t

c. Encrypt update: $C_i = Enc_{K_i}(\Delta w_i^t)$

d. Compute HMAC: MAC_i

e. Send $(C_i, MAC_i)$to server

Server Side:

f. Verify authenticity via HMAC

g. Decrypt valid updates

h. Aggregate updates → compute w^{t+1}

i. Broadcast $w^{t+1}$to all clients

Where N is the number of the clients involved, S the size of each update in bits to the model, and R the count of the communication rounds per unit time. There is a certain amount of secret key material used by each client per round, denoted as K_c whereas the QKD system offers keys with a rate of K_q bits per second. In this architecture, model updates are not directly encrypted with QKD, but it provides symmetric keys that are encrypted with the Advanced Encryption Standard (AES) to achieve confidentiality and authenticated with HMAC to achieve integrity and authentication. Therefore, the overall consumption of keys per round varies linearly with the number of clients, and can be represented as N × Kc. This formulation allows

assessing the rate of the QKD key generation rate K_q is enough to support the safe running of the federated learning process as the system scale and frequency of communication grow. The key consumption per round is given by:

$$K_{total} = N.K_C \quad (5)$$

The model shows that QKD overhead is independent of model update size, since symmetric encryption (e.g., AES-256) is used for bulk data protection. Instead, QKD overhead scales with the number of clients and round frequency through key consumption.

4. Results and Discussions

The experimental study aims to evaluate (i) the impact of integrating QKD-derived symmetric keys on the federated learning (FL) training process in terms of convergence and overhead, and (ii) the robustness of the proposed secure update exchange pipeline to poisoning attacks when robust aggregation is employed. Due to the research focused on security and protocol-level costs (rather than optimizing a particular ML benchmark), we conducted a controlled simulation that models FL behavior with and without QKD-based security. The simulator abstracts local training convergence curves and models communication and computation overheads introduced by cryptographic primitives. Table 2 illustrates the key Simulation Parameters Used in QKD-Enabled Federated Learning Experiments.

To assess the effectiveness of the proposed QKD-enabled secure model update exchange framework, four experimental scenarios were evaluated. All the above scenarios include both cooperative and competing FL settings, allowing in-depth analysis of various scenarios in terms of robustness and efficiency.

The experimental evaluation uses a simulator that models the interaction between federated learning workflows and quantum-secured communication. Rather than focusing on a specific deep learning architecture [24, 25], the simulator represents local training behavior using convergence profiles derived from standard federated learning benchmarks.

This abstraction allows system-level performance metrics, including communication latency, cryptographic overhead, and key-management cost, to be analyzed in a controlled setting, which is the primary focus of this study. Secure communication is modeled using the BB84 protocol, including realistic key-generation delays, authentication overhead, and periodic key-refresh operations. The simulator incorporates both quantum and classical channel interactions together with the costs of message authentication and encryption. In each communication round, the overall client latency is approximated as the sum of local training time, encryption and authentication overhead, network transmission latency, and any additional cost associated with QKD key generation or key refresh, where applicable. On the server side, latency also includes the time required to verify the integrity of received updates and perform aggregation. These components are combined within the simulator to estimate both per-round completion time and end-to-end training time. To study performance–security trade-offs, we consider three main scenarios: a baseline federated learning setting without additional security, in which updates are transmitted in plaintext; a QKD-secured federated learning setting, in which updates are encrypted with QKD-generated keys and authenticated using HMAC while retaining the same aggregation strategy; and adversarial settings, in which compromised clients are introduced and system behavior is evaluated using conventional mean aggregation and robust aggregation methods such as median and trimmed mean.

Table 2. Simulation Parameters

Parameter	Value
Number of clients	20
Communication rounds	50
Model update size (before compression)	2 MB per client per round
QKD classical-tag overhead	64 B (MAC) + 512 B protocol metadata per client per round
Local training time per round (per client)	6.0 s
Transmission time per round (per client)	0.2 s
QKD key generation cost (per client, per session)	1.5 s (amortized or periodic)
Encryption + MAC computation time per round (per client)	0.05 s
Server verification & aggregation time per round	0.1 s

Figure 2 illustrates the accuracy of the simulated global model over 50 rounds for the four experimental scenarios. The QKD-secured FL curve closely follows the baseline, suggesting that, when correctly implemented, encryption and authentication do not hinder model convergence. Under poisoning attacks, the accuracy of simple mean aggregation drops substantially compared with the baseline. However, much of this loss is recovered when robust aggregation methods such as median or trimmed mean are used. This result highlights the importance of combining authentication with robust aggregation to mitigate poisoning attacks.

Figure 3 compares the total number of bytes transmitted per round. In the baseline setting, only the 2 MB model updates are sent by each client. In the QKD-secured setting, per-client overhead from the MAC and protocol header increases total communication only slightly, by approximately 0.011 MB per client in our simulation. With aggressive 4× compression, the payload can be reduced substantially while still preserving authentication. The simulation therefore shows that the additional communication cost introduced by QKD-related classical metadata is small relative to the model payload, and that compression can make the overall security overhead more cost-effective.

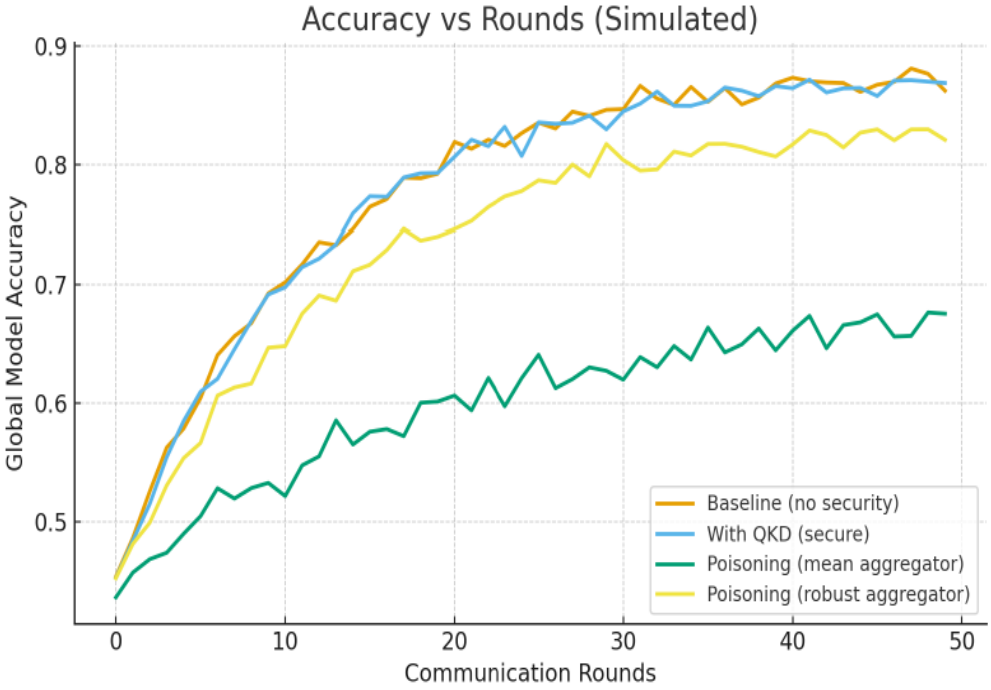

Figure 2. Accuracy vs Rounds

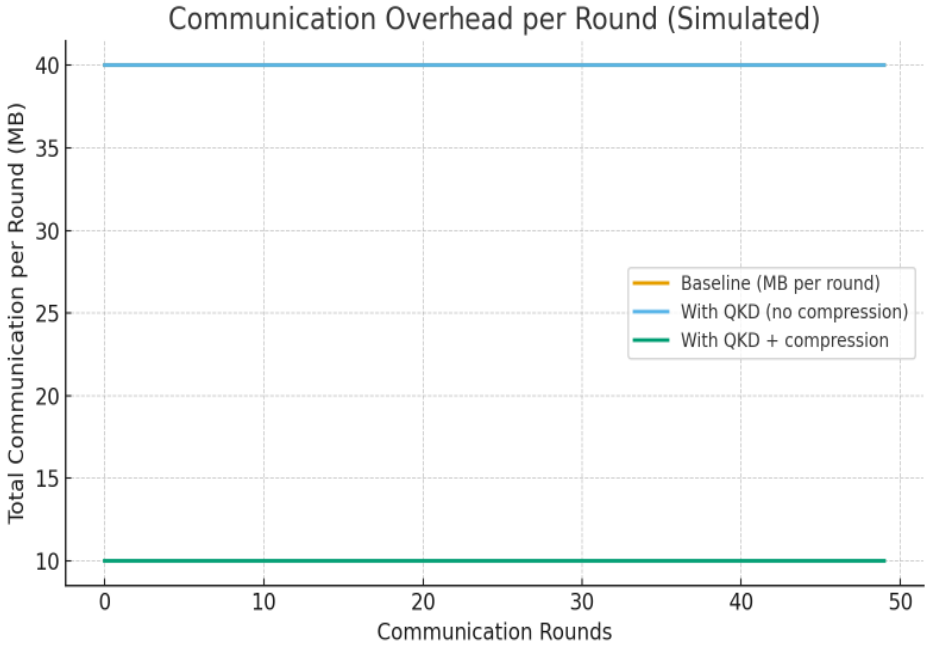

Figure 3. Communication overhead per round

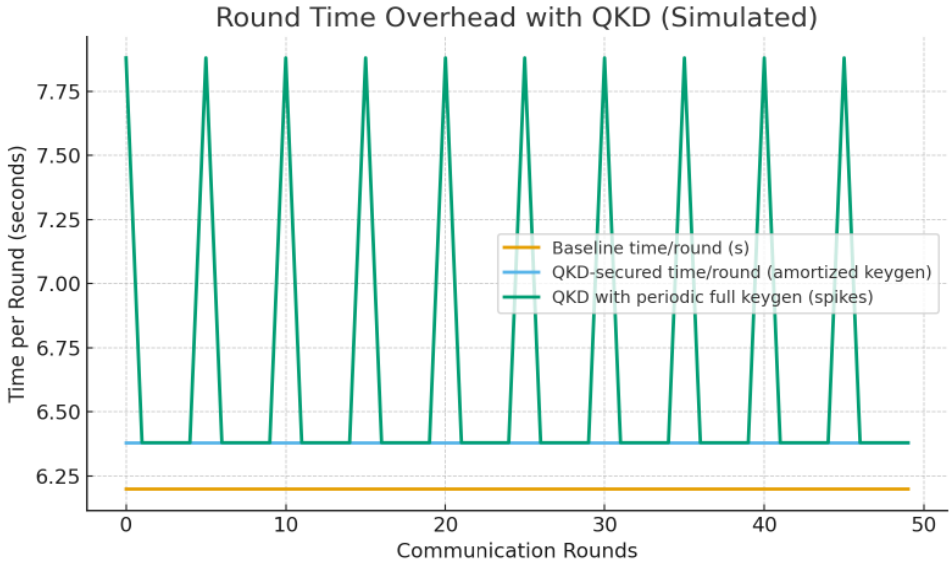

Figure 4. Time per round (QKD overhead)

Table 3. Summary of Accuracy, Communication Cost, and Time per Round

Scenario	Final Accuracy	Avg Communication / Round	Avg Time / Round
Baseline FL	0.862	40 MB	6.2 s
QKD-Secured FL	0.869	40.01 MB	6.38 s
QKD-Secured + Compression	–	10.01 MB	–
Poisoning (Mean Aggregation)	0.675	–	–
Poisoning (Robust Aggregation)	0.821	–	–
QKD (Periodic Key Generation)	–	–	6.68 s

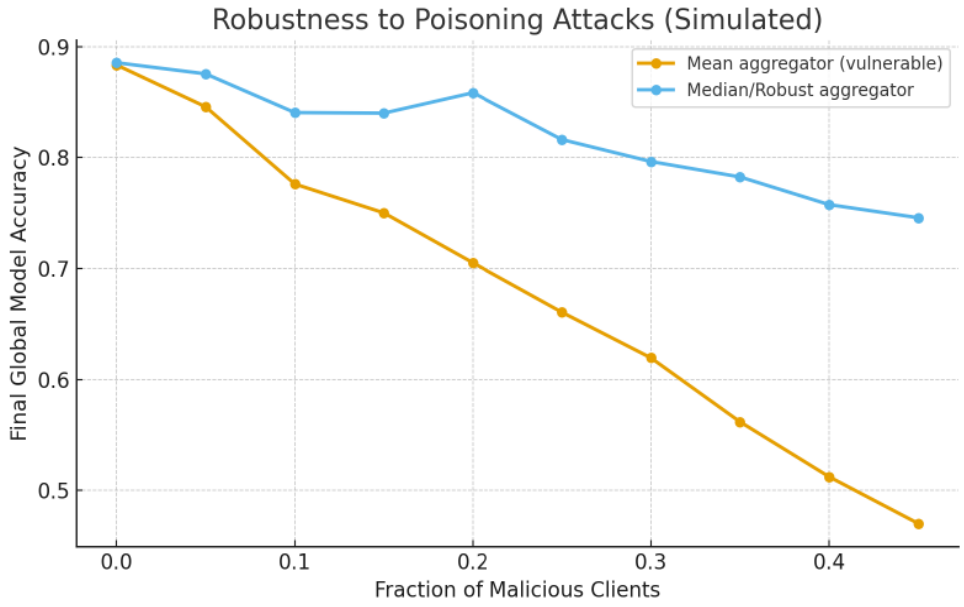

Figure 5. Robustness vs malicious fraction

Table 4. Comparative Analysis

Ref.No.	Work	Security Technique	Target Environment	Key Outcomes Reported	Limitations	Comparison with Proposed QKD-FL Model
[6]	Multi-layered framework for post-quantum secure machine learning integrating cryptographic protections	Post-Quantum Cryptography (lattice-based encryption & signatures)	Edge and Cloud-based distributed machine learning systems	Provides end-to-end ML security with quantum-resistant encryption	High computational cost; lacks real-world large-scale deployment validation	QKD-FL model offers stronger information-theoretic key security via QKD
[14]	A post-quantum secure aggregation for federated learning	Post-Quantum Secure Aggregation	Federated Learning	Improved resistance against quantum attacks during aggregation	Higher computational overhead and no real-time key exchange	Proposed model enhances security using dynamic QKD-based secret key exchange with lower attack exposure and stronger confidentiality
[16]	PQSF: Post-quantum secure privacy-preserving federated learning	Post-Quantum Cryptography	Privacy-Preserving FL	Enhanced privacy preservation against future quantum adversaries	Key management complexity and increased encryption cost	Proposed framework introduces QKD-based automatic symmetric key generation, reducing key distribution risks and improving secure model update exchange
[17]	Post-Quantum Cryptography Blockchain-based Federated Learning	Post-Quantum Cryptography	Federated Learning Networks	Quantum-safe and traceable FL framework	Blockchain overhead and scalability challenges	QKD-FL provides quantum-secure key distribution with lower blockchain dependency and reduced computational overhead
[24]	Lightweight Post-Quantum Cryptography	Lightweight PQC	IoT, Blockchain, E-Learning	Quantum-resistant security	Computational overhead, large keys	QKD-FL offers secure quantum-safe key exchange and dynamic key management

						for federated learning
Proposed Model	QKD-enabled Secure Model Update Exchange in Federated Learning	Quantum Key Distribution + Secure Federated Learning	Distributed Federated Learning Systems	Enhanced confidentiality, quantum-resistant secure model exchange, eavesdropping detection, reduced key compromise risk, and secure aggregation support	Requires QKD infrastructure deployment	Outperforms existing methods in quantum-resilient secure model update transmission and dynamic secret key generation

Figure 4 visualizes the wall-clock time per round. In the baseline setting, local training (6 s) and transmission (0.2 s) dominate the per-round runtime. In the QKD-secured setting, the additional time arises from amortized key generation, when keys are generated periodically during a session, together with small per-round encryption and verification costs. Under steady-state operation, amortization keeps the average per-round overhead close to zero. Periodic key generation, for example full key regeneration every five rounds, can introduce measurable spikes in round time, but these remain acceptable given the assumed QKD link rates. This indicates that the key-renewal frequency should be selected according to QKD link capacity and the latency tolerance of the target deployment.

Figure 5 plots final accuracy as a function of the fraction of malicious clients, comparing mean aggregation with median-based robust aggregation. The final accuracy of the mean aggregator decreases sharply as the malicious fraction increases. In contrast, the median-based robust aggregator shows substantially less degradation, indicating that the combination of QKD-backed authentication, which mitigates man-in-the-middle tampering, and robust aggregation provides stronger protection against poisoning attacks. Table 3 summarizes the key performance measures observed across the different experimental settings, including final model accuracy, communication cost per round, and time per round in both secure and unsecured environments.

Final model accuracy and average time per round are not reported because the simulation primarily focused on communication reduction effects, rather than detailed convergence behavior in this preliminary analysis. Average communication per round and average time per round are not reported, as these scenarios were simulated specifically to evaluate learning-layer robustness under adversarial updates, not communication-layer overhead. Communication and time measures would also be comparable to the baseline FL or QKD-secured FL runs, so their reporting would not be informative. Final model accuracy and communication per round are not reported because this experiment isolates the impact of periodic QKD key refreshes on system latency, without changing the learning or transmission behavior. The focus is on the time overhead introduced by key generation, which is reported. In the proposed framework, security mechanisms enabled via the BB84 protocol are integrated strictly at the communication stage. Specifically, each client performs local model training independently, after which the computed model updates are encrypted and authenticated prior to transmission to the central server. Key generation and refresh operations occur periodically and are decoupled from the gradient computation process. As a result, these security operations do not modify the optimization procedure, loss function, or gradient updates, but instead introduce additional latency in the communication pipeline. The results indicate that the inclusion of QKD-based secure communication does not alter the convergence trajectory of the federated learning process. The number of communication rounds required to reach a target accuracy remains consistent with the baseline system. Therefore, the term “negligible convergence impact” refers to the preservation of training dynamics rather than identical point-wise accuracy values. The experimental findings demonstrate that QKD-enabled secure communication preserves the functional performance of federated learning while introducing modest and predictable overhead in communication latency. This confirms that strong communication security can be achieved without degrading model convergence. In general, the table indicates that FL secured by QKD has good security with little effects on accuracy as well as low overhead, whereas good aggregation resistance to poisoning attacks is demonstrated by robust aggregation. Compression schemes also reduce communication overheads and this framework is even possible in resource constrained federated networks.

Table 4 compares the proposed QKD-enabled federated learning framework with recent state-of-the-art secure federated learning approaches in terms of security mechanism, deployment environment, outcomes, and limitations. The comparative analysis demonstrates that existing approaches mainly rely on post-quantum cryptographic algorithms or blockchain-assisted protection mechanisms, which often introduce computational complexity, latency, or key management challenges. In contrast, the proposed QKD-enabled framework provides dynamic quantum-secure key exchange with inherent eavesdropping detection, thereby strengthening the confidentiality and integrity of federated model updates against future quantum-enabled attacks.

5. Conclusion

This work proposes a Quantum Key Distribution (QKD)-based secure model-update exchange framework for Federated Learning (FL) to address privacy and integrity risks in collaborative model training. The proposed architecture combines the BB84 QKD protocol with classical cryptographic mechanisms such as AES and HMAC to protect model updates during transmission. A system-level simulation framework is used to evaluate the interaction between federated learning communication and QKD-based key management. The results indicate that strong communication security can be achieved with limited and predictable overhead, while preserving the convergence behavior of the learning process. Because the security mechanisms operate at the communication layer, they do not directly improve model accuracy or address learning-layer threats such as model poisoning or malicious clients; these issues require additional defensive mechanisms. Although the present simulation provides useful insight into feasibility and scalability, further validation using real federated learning benchmarks and deployment-oriented testbeds remains an important direction for future work. Integrating robust aggregation and advanced adversarial defenses with the proposed secure communication framework would be a promising next step toward end-to-end security. Overall, this study provides a practical basis for integrating quantum-secured communication into federated learning systems while clearly identifying both its potential benefits and its current constraints.

References

[1] V. Mothukuri, R.M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, G. Srivastava, A Survey on Security and Privacy of Federated Learning. Future Generation Computer Systems, 115, (2021) 619–640. https://doi.org/10.1016/j.future.2020.10.007

[2] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H.B. McMahan, S. Patel, D. Ramage, A. Segal, K. Seth, Practical Secure Aggregation for Privacy-Preserving Machine Learning. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS), (2017) 1175–1191. https://doi.org/10.1145/3133956.3133982

[3] H.U. Manzoor, A. Shabbir, A. Chen, D. Flynn, A. Zoha, A Survey of Security Strategies in Federated Learning: Defending Models, Data, and Privacy. Future Internet, 16(10), (2024) 374. https://doi.org/10.3390/fi16100374

[4] B. Liu, N. Lv, Y. Guo, Y. Li, Recent Advances on Federated Learning: A Systematic Survey. Neurocomputing, 597, (2024) 128019. https://doi.org/10.1016/j.neucom.2024.128019

[5] L. Cui, Y. Qu, G. Xie, D. Zeng, R. Li, S. Shen, S. Yu, Security and Privacy-enhanced Federated Learning for Anomaly Detection in IoT Infrastructures. IEEE Transactions on Industrial Informatics, IEEE, 18(5), (2022) 3492–3500. https://doi.org/10.1109/TII.2021.3107783

[6] R. Hamza, A. Alotaibi, K. Muhammad, A practical multi-layered framework for post-quantum secure machine learning, Engineering Applications of Artificial Intelligence, 163, (2026) 113044. https://doi.org/10.1016/j.engappai.2025.113044

[7] L. Chen, Q. Chen, M. Zhao, J. Chen, S. Liu, Y. Zhao, DDKA-QKDN: Dynamic On-Demand Key Allocation Scheme for Quantum Internet of Things Secured by QKD Network. Entropy, 24, (2022),149. https://doi.org/10.3390/e24020149

[8] X. Qin, R. Xu, Efficient Post-Quantum Cross-Silo Federated Learning based on Key Homomorphic Pseudo-Random Function. Mathematics, 13(9), (2025) 1404. https://doi.org/10.3390/math13091404

[9] Y. Xu, T. Wang, P. Huang, G. Zeng, Integrated Distributed Sensing and Quantum Communication Networks. Research, 7, (2024) 0416. https://doi.org/10.34133/research.0416

[10] H. Gharavi, E. Monteiro, J. Granjal, PQBFL: A Post-Quantum Blockchain-Based Protocol for Federated Learning. Computer Networks, 269, (2025) 111472. https://doi.org/10.1016/j.comnet.2025.111472

[11] B. Appiah, I. Osei, B.K. Frimpong, D. Commey, K. Owusu-Agymang, G. Assamah, Enhanced Federated Learning for Secure Medical Data Collaboration. Journal of Analytical Science and Technology, 16, (2025) 13. https://doi.org/10.1186/s40543-025-00484-2

[12] M. Zahidy, D. Ribezzo, C. De Lazzari, I. Vagniluca, N. Biagi, R. Müller, T. Occhipinti, L.K. Oxenløwe, M. Galili, T. Hayashi, D. Cassioli, Practical High-Dimensional Quantum Key Distribution Protocol over Deployed Multicore Fiber. Nature Communications, 15(1), (2024) 1651. https://doi.org/10.1038/s41467-024-45876-x

[13] P. Kairouz, H.B. McMahan, Advances and Open Problems in Federated Learning. Foundations and Trends in Machine Learning, 14(1–2), (2021) 1–210. https://doi.org/10.1561/2200000083

[14] S. Yang, Y. Chen, S. Tu, Z. Yang, A Post-Quantum Secure Aggregation for Federated Learning. Proceedings of the 2022 12th International Conference on Communication and Network Security (ICCNS '22), (2023) 117–124. https://doi.org/10.1145/3586102.3586120

[15] Y. Feng, Y. Guo, Y. Hou, Y. Wu, M. Lao, T. Yu, G. Liu, A Survey of Security Threats in Federated Learning. Complex & Intelligent Systems, 11(2), (2025) 165. https://doi.org/10.1007/s40747-024-01664-0

[16] X. Zhang, H. Deng, R. Wu, J. Ren, Y. Ren, PQSF: Post-Quantum Secure Privacy-Preserving Federated Learning. Scientific Reports, 14, (2024) 23553. https://doi.org/10.1038/s41598-024-74377-6

[17] D. Commey, G.V. Crosby, PQS-BFL: A post-quantum secure blockchain-based federated learning framework. Expert Systems with Applications, 312, (2026) 131449. https://doi.org/10.1016/j.eswa.2026.131449

[18] M. Bilal Yaseen, F. Wan, S. Asif, J. Wan, QFL-SecEdge: A quantum federated learning framework with post-quantum cryptographic task distribution for ultra-reliable mission-critical 6G-IoT networks, IEEE Internet of Things Journal, 2026. https://doi.org/10.1109/JIOT.2026.3685072

[19] J. Lai, F. Yao, J. Wang, M. Zhang, F. Li, W. Zhao, H. Zhang, Application and Development of QKD-based Quantum Secure Communication. Entropy, 25(4), (2023) 627. https://doi.org/10.3390/e25040627

[20] A. Gupta, M. Kumar Maurya, K. Dhere, V. Kumar Chaurasiya, "Privacy-Preserving Hybrid Federated Learning Framework for Mental Healthcare Applications: Clustered and Quantum Approaches," in IEEE Access, vol. 12, pp. 145054-145068, 2024. https://doi.org/10.1109/ACCESS.2024.3464240

[21] V. Zapatero, T. Van Leent, R. Arnon-Friedman, W.Z. Liu, Q. Zhang, H. Weinfurter, M. Curty, Advances in Device-Independent Quantum Key Distribution. Npj Quantum Information, 9(1), (2023) 10. https://doi.org/10.1038/s41534-023-00684-x

[22] S. K. Shandilya, C. Ganguli, A. Kumar, I. Izonin and M. Gregus, Post-Quantum Cryptography and Nature-Inspired Cyber Defense: Strategic Readiness and Adaptive Techniques for Next-Gen Threat Response. IEEE Access, 14, (2026) 27418-27434. http://doi.org/10.1109/ACCESS.2026.3663832

[23] R. Rahman, S. Shaham and D. C. Nguyen, "Toward Personalized Quantum Federated Learning for Anomaly Detection," in IEEE Transactions on Network Science and Engineering, vol. 13, pp. 3335-3350, 2026. https://doi.org/10.1109/TNSE.2026.3654089

[24] C. L. Chen, K. W. Zeng, W. Y. Li, C. F. Lee, L. C. Liu, Y. Y. Deng, Lightweight Post-Quantum Cryptography: Applications and Countermeasures in Internet of Things, Blockchain, and E-Learning. Engineering Proceedings, 103(1), (2025) 14. https://doi.org/10.3390/engproc2025103014

[25] D. Javeed, M.S. Saeed, I. Ahmad, M. Adil, P. Kumar, A.N. Islam, Quantum-Empowered Federated Learning and 6G Wireless Networks for IoT Security: Concept, Challenges and Future Directions. Future Generation Computer Systems, 160, (2024) 577–597. https://doi.org/10.1016/j.future.2024.06.023

Authors Contribution Statement

Venkadeshan Ramalingam: Conceptualization, Methodology, Data collection, Formal analysis, Writing - Original Draft. Basant Kumar: Conceptualization, Validation, Supervision, Writing - Original Draft. S. Jagadeesan: Conceptualization, Methodology, Writing - Review & Editing. T. Vetriselvi: Writing - Review & Editing. T. Sivakumar: Writing - Review & Editing. All the authors read and approved the final version this manuscript.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Has this article screened for similarity?

Yes

Data Availability

The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

About the License