Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

A Survey on AI-Driven Threat Detection in Quantum Key Distribution Systems with Automated Response Mechanisms

Prof. Mohini A. Thorat[1], Janhvi Chavan[2], Anushka Kumbhar[3], Mithilesh Kadam[4], Aditya Shinde[5]
[1,2,3,4,5]Department of Computer Engineering, JSPM's Jayawantrao Sawant College of Engineering (JSCOE), Pune, India

***Abstract*—Quantum computing advancements pose a growing threat to conventional cryptographic systems based on computational hardness assumptions such as RSA and Elliptic Curve Cryptography. Quantum Key Distribution (QKD) has emerged as a theoretically secure alternative by leveraging the fundamental principles of quantum mechanics to establish shared secret keys between communicating parties. However, practical QKD deployments remain vulnerable to channel noise, hardware imperfections, side-channel attacks, and sophisticated intercept-resend strategies that exploit implementation weaknesses. Traditional QKD security analysis relies on static Quantum Bit Error Rate (QBER) thresholds, which are insufficient for distinguishing malicious eavesdropping from natural environmental disturbances. Recent advancements in artificial intelligence and machine learning have introduced intelligent anomaly detection techniques capable of analyzing complex quantum channel behaviour in real time. This paper presents a comprehensive survey of AI-driven threat detection and response systems for quantum cryptographic networks. Various techniques including Random Forest classifiers, deep learning models, Support Vector Machines, and reinforcement learning-based adaptive response mechanisms are discussed along with their strengths and limitations. The paper also highlights major research challenges such as false-positive detection, computational overhead, hardware dependency, and limited integration of automated mitigation strategies. Furthermore, a conceptual multimodal framework combining IBM Qiskit-based BB84 simulation, machine learning-based threat classification, and automated Privacy Amplification is discussed as a promising solution for securing next-generation quantum communication infrastructures.**

***Index Terms*—Quantum Key Distribution, QKD, BB84, quantum cryptography, AI-driven threat detection, anomaly detection, Random Forest, machine learning, QBER, automated mitigation, quantum security.**

I. INTRODUCTION

Quantum computing is rapidly evolving from a theoretical concept to a practical engineering challenge, raising fundamental questions about the long-term security of modern communication infrastructure. Cryptographic schemes widely deployed in banking, government, and defence sectors — including RSA, Diffie-Hellman, and Elliptic Curve Cryptography — derive their security from the computational difficulty of factoring large integers and solving discrete logarithm problems. However, Shor's quantum algorithm has demonstrated the theoretical capability to solve these problems exponentially faster than classical computers, potentially rendering existing public-key infrastructure obsolete within the foreseeable future [3].

Quantum Key Distribution (QKD) has been proposed as a fundamental solution to this threat. Unlike classical key exchange protocols, QKD derives its security from the laws of quantum mechanics rather than computational assumptions. The BB84 protocol, introduced by Bennett and Brassard, enables two parties to establish a shared secret key while guaranteeing that any eavesdropping attempt will introduce measurable disturbances into the quantum channel due to the No-Cloning Theorem and wave-function collapse [1].

Although QKD offers information-theoretic security in theory, practical deployments face several significant challenges. Environmental noise, photon transmission losses, detector imperfections, and side-channel vulnerabilities can degrade channel fidelity and open windows for sophisticated attacks. Among these, intercept-resend attacks, photon number splitting attacks, and Trojan horse attacks remain serious concerns for real-world QKD systems [5].

Traditional QKD security monitoring relies on static Quantum Bit Error Rate (QBER) thresholds to detect potential eavesdropping. However, this approach is inadequate for distinguishing malicious interception from legitimate environmental noise, resulting in both false positives and undetected attacks. Recent research has demonstrated that artificial intelligence and machine learning techniques can significantly improve QKD threat detection by learning complex patterns in quantum channel behaviour and enabling dynamic, context-aware security decisions [7], [8].

This paper surveys recent AI-driven threat detection and automated response systems proposed for quantum cryptographic networks, discusses their strengths and limitations, and explores integrated frameworks that combine quantum simulation with intelligent security operations.

II. BACKGROUND

Artificial Intelligence and Machine Learning technologies have become essential tools in modern cybersecurity and network anomaly detection systems. In the context of quantum cryptography, these techniques are being applied to analyse quantum channel metrics and identify deviation patterns indicative of malicious activity.

Quantum Key Distribution protocols, particularly BB84 and E91, form the cryptographic foundation of quantum-secure communication. BB84 uses four polarization states encoded in two conjugate bases to transmit key bits, while E91 utilises

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

quantum entanglement and Bell inequality violations to detect eavesdropping [2]. Both protocols rely on QBER analysis to assess channel security. High QBER values indicate either severe noise or active eavesdropping.

IBM Qiskit is an open-source quantum computing framework that enables high-fidelity simulation of quantum circuits and communication channels. AerSimulator, part of the Qiskit ecosystem, supports realistic noise modelling including depolarizing noise, thermal relaxation, and gate errors, making it suitable for generating representative quantum communication datasets for machine learning training [9].

Machine learning classifiers such as Random Forest, Support Vector Machines (SVM), and Neural Networks have demonstrated strong performance in network intrusion detection tasks due to their ability to model non-linear relationships in high-dimensional feature spaces [7]. In QKD systems, features such as QBER, noise level, sifted key length, effective key rate, and engineered heuristic metrics can be extracted per transmission session and used to train supervised classifiers for binary threat classification.

Automated mitigation mechanisms including Privacy Amplification, Information Reconciliation, and protocol migration (e.g., switching from BB84 to E91) represent the response layer of intelligent QKD security frameworks. These mechanisms reduce adversarial knowledge and restore channel security following detected threats [6].

Security Operations Center (SOC) dashboards integrated with real-time quantum channel monitoring provide human-readable situational awareness and trigger automated countermeasures when anomalies are detected, enabling scalable quantum network security management.

III. LITERATURE SURVEY

Several researchers have proposed AI-based threat detection systems for quantum cryptographic networks using machine learning and quantum simulation techniques.

Scarani et al. [5] presented a comprehensive review of the security of practical QKD systems, analysing vulnerabilities arising from device imperfections, side-channel leakage, and implementation flaws. Their work established the theoretical basis for QBER-based security thresholds and highlighted the limitations of static threshold approaches in noisy practical channels.

Pirandola et al. [6] conducted a detailed survey of advances in quantum cryptography, covering discrete-variable and continuous-variable QKD, device-independent protocols, and satellite-based quantum networks. The study discussed deployment challenges including atmospheric noise, distance limitations, and the integration of quantum repeaters, underscoring the need for adaptive and intelligent security monitoring systems.

Breiman [7] introduced the Random Forest ensemble learning algorithm, demonstrating its superior performance in classification tasks with high-dimensional feature spaces and resistance to overfitting. Random Forest has since been widely adopted in network intrusion detection and anomaly classification due to its accuracy, interpretability, and computational efficiency.

Sommer and Paxson [8] critically evaluated the application of machine learning to network intrusion detection, identifying key challenges including the distribution shift between training data and real traffic, the high cost of false positives in operational systems, and the interpretability of black-box models. Their findings are directly relevant to AI-driven QKD threat detection design.

Liao et al. [11] reviewed intrusion detection systems across classical networks and identified Random Forest and SVM-based approaches as consistently high-performing models for binary and multiclass attack classification. Their comparative analysis informed the selection of classifiers suitable for quantum channel anomaly detection.

Recent studies have also explored deep learning architectures including Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) for temporal anomaly detection in quantum channel sequences [13]. These models capture sequential dependencies in QBER time-series data, enabling detection of slow, stealthy eavesdropping attacks that evade static threshold systems.

Additionally, reinforcement learning-based adaptive mitigation frameworks have been proposed for autonomous quantum network security management [12]. These systems dynamically adjust Privacy Amplification parameters, basis reconciliation strategies, and protocol selections in response to detected threat levels, reducing human intervention requirements in large-scale quantum networks.

IV. PROBLEM FORMULATION

Despite significant advances in both quantum cryptography and machine learning, several challenges continue to limit the reliability and scalability of AI-driven QKD threat detection systems. Existing approaches often rely on isolated QBER thresholds or single-modality anomaly detection without integrating environmental context or adaptive response mechanisms.

Vision-independent AI classifiers trained on simulated datasets may produce false-positive alerts when deployed in real quantum channels where noise characteristics differ from training distributions. Simulation environments such as IBM Qiskit provide controlled but idealised noise models that may not fully capture hardware-specific imperfections present in photonic QKD implementations.

Most current AI-based QKD security systems also lack automated mitigation layers. Detection of a threat typically triggers a manual response rather than an immediate automated countermeasure, introducing response delays that can compromise key security in time-sensitive applications.

Another significant challenge is the high computational demand of deep learning architectures when applied to real-time quantum channel monitoring. LSTM and transformer-based models require substantial memory and processing resources

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

that may be unavailable in embedded or resource-constrained quantum network nodes.

Table I summarises the key research gaps identified in existing systems and their impact on practical QKD security deployments.

Research Gap	Impact on Existing Systems
Static QBER thresholds	Cannot distinguish malicious attacks from environmental noise
Simulation-to-hardware gap	Reduces classifier accuracy in real deployments
Lack of automated mitigation	Increases response latency during active attacks
High computational cost	Limits real-time deployment on edge QKD nodes
Single-modality detection	Decreases contextual awareness and detection accuracy
Limited feature engineering	Reduces sensitivity to stealthy low-QBER attacks

TABLE I
RESEARCH GAPS AND THEIR IMPACT ON QKD SECURITY SYSTEMS

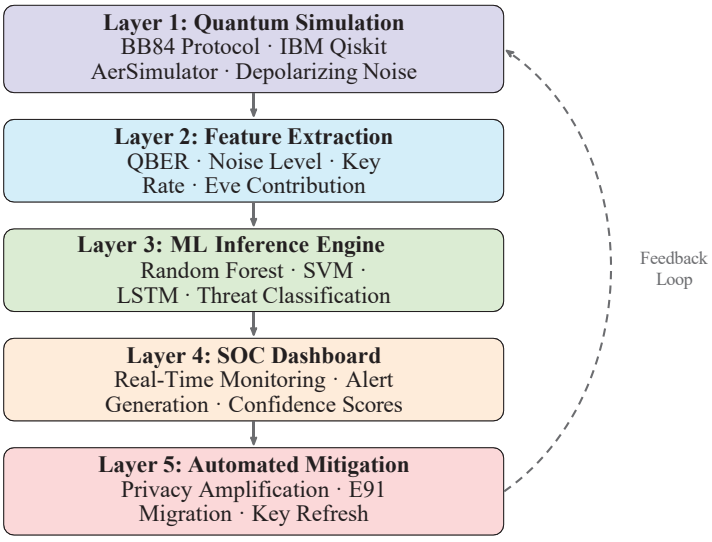

Fig. 1. Proposed five-layer AI-driven QKD threat detection and automated response framework.

These limitations highlight the need for intelligent multimodal QKD security frameworks that combine quantum simulation, AI-based anomaly detection, and automated response mechanisms within a unified architecture.

V. SOLUTION DOMAIN

To overcome the limitations of traditional QKD security monitoring, multimodal frameworks integrating quantum simulation, machine learning classifiers, and automated response mechanisms have emerged as promising solutions. A conceptual framework combining IBM Qiskit-based BB84 simulation with AI-driven threat detection represents one such approach.

The proposed solution domain centres on a five-layer architecture: Quantum Simulation, Feature Extraction, Machine Learning Inference, Security Operations Center (SOC) Monitoring, and Automated Mitigation, as illustrated in Fig. 1.

The Quantum Simulation Layer generates realistic BB84 communication data using IBM Qiskit's AerSimulator with configurable depolarizing noise, enabling high-fidelity training datasets that represent both secure and compromised channel conditions [9].

Feature extraction transforms raw quantum communication records into structured security metrics including QBER, noise level, sifted key length, effective key rate, and an engineered Eve Contribution feature that isolates attacker-induced error from natural channel depolarization. This heuristic feature significantly reduces false-positive rates by removing expected environmental noise contributions from the anomaly signal [5].

Machine learning classifiers, particularly Random Forest models trained on large simulated datasets, provide accurate and efficient binary classification of quantum channel sessions as secure or compromised. Random Forest is particularly suitable for this task due to its resistance to overfitting, ability to rank feature importance, and low inference latency compared with deep learning alternatives [7].

The SOC dashboard provides real-time visualisation of channel security metrics and classifier confidence scores, enabling security operators to monitor quantum network health. Automated mitigation mechanisms including Privacy Amplification and migration to the E91 entanglement-based protocol are triggered automatically upon high-confidence threat detection, restoring communication security without manual intervention [2].

Such integrated architectures provide scalable and reliable solutions for real-time quantum network security management across enterprise quantum communications, financial transaction security, government data protection, and critical infrastructure defence applications.

VI. CONCLUSION

Artificial intelligence, machine learning, and quantum cryptography together represent a transformative opportunity for securing future communication infrastructure against both classical and quantum adversaries. Existing research demonstrates that AI-based frameworks significantly outperform static QBER threshold methods in accurately detecting malicious interception attempts in Quantum Key Distribution networks. However, challenges including simulation-to-hardware generalisation, computational constraints, and the absence of automated response layers continue to limit practical deployment.

This survey reviewed various AI-driven QKD threat detection techniques including Random Forest classifiers, deep learning temporal models, and reinforcement learning-based adaptive mitigation systems. The study also discussed current research gaps including static threshold limitations, single-modality detection, and high computational overhead, and highlighted the growing importance of integrating automated

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 06 , June - 2026

countermeasures with AI-based anomaly detection for end-to-end quantum security.

Overall, multimodal frameworks combining quantum simulation, intelligent anomaly detection, and automated mitigation represent a promising direction for building scalable, reliable, and context-aware quantum network security systems capable of protecting next-generation quantum communication infrastructures against sophisticated adversarial threats.

REFERENCES

[1] C. H. Bennett and G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing," *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, Bangalore, India, 1984.
[2] A. K. Ekert, "Quantum Cryptography Based on Bell's Theorem," *Physical Review Letters*, vol. 67, no. 6, pp. 661–663, 1991.
[3] P. W. Shor, "Algorithms for Quantum Computation: Discrete Logarithms and Factoring," *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, 1994.
[4] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum Cryptography," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 145–195, 2002.
[5] V. Scarani et al., "The Security of Practical Quantum Key Distribution," *Reviews of Modern Physics*, vol. 81, no. 3, pp. 1301–1350, 2009.
[6] S. Pirandola et al., "Advances in Quantum Cryptography," *Advances in Optics and Photonics*, vol. 12, no. 4, pp. 1012–1236, 2020.
[7] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[8] R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," *IEEE Symposium on Security and Privacy*, 2010.
[9] H. Abraham et al., "Qiskit: An Open-Source Framework for Quantum Computing," IBM Research, 2024.
[10] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2010.
[11] H. J. Liao, C. H. R. Lin, Y. C. Lin, and K. Y. Tung, "Intrusion Detection System: A Comprehensive Review," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 16–24, 2013.
[12] S. Garcia, M. Grill, J. Stiborek, and A. Zunino, "An Empirical Comparison of Botnet Detection Methods," *Computers & Security*, vol. 45, pp. 100–123, 2014.
[13] X. Zhao, Y. Li, and J. Zhang, "Deep Learning-Based Anomaly Detection for Quantum Channel Security Monitoring," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 2134–2146, 2022.
[14] H.-K. Lo, M. Curty, and K. Tamaki, "Secure Quantum Key Distribution," *Nature Photonics*, vol. 8, pp. 595–604, 2014.
[15] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, "Overcoming the Rate-Distance Limit of Quantum Key Distribution Without Quantum Repeaters," *Nature*, vol. 557, pp. 400–403, 2018.
[16] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, "Secure Quantum Key Distribution with Realistic Devices," *Reviews of Modern Physics*, vol. 92, no. 2, p. 025002, 2020.
[17] R. Alle´aume et al., "Using Quantum Key Distribution for Cryptographic Purposes: A Survey," *Theoretical Computer Science*, vol. 560, pp. 62–81, 2014.
[18] I. Vagniluca et al., "Efficient Time-Bin Encoding for Practical High-Dimensional Quantum Key Distribution," *Physical Review Applied*, vol. 14, p. 014051, 2020.