

AI-Driven Approaches to Post-Quantum Cybersecurity: Bridging Quantum Computing Threats and Intelligent Defense Systems

Rahul Gandhi

Liverpool Business School
Email: [gandhirahul230\[at\]gmail.com](mailto:gandhirahul230[at]gmail.com)

Abstract: *Quantum computers will eventually break the encryption schemes protecting most of today's digital infrastructure. Because adversaries can harvest encrypted data now and decrypt it later once that capability exists, organizations face an immediate prioritization problem: which data should be migrated to post-quantum cryptography first? We propose a Quantum Exposure Window (QEW) metric, grounded in published expert forecasts of quantum arrival timelines, combined with a machine-learning anomaly detector that flags likely data-harvesting activity in network telemetry. We tested the resulting scheduling algorithm in a controlled simulation and found that our original design- combining these two signals multiplicatively - failed to prioritize short-lifespan data under active attack. We identify why this happens, derive a corrected additive rule, and prove it eliminates the failure mode; the simulation confirms an improvement from 0% to 75.6% coverage of the affected cases at the same migration budget. We discuss the limits of this validation, including its synthetic nature and an unresolved assumption about linking network anomalies to specific data assets, and outline the real-world testing this would require before deployment.*

Keywords: Post-Quantum Cryptography; Artificial Intelligence; Cybersecurity; Threat Detection; Quantum Computing

1. Introduction

Quantum computers capable of running Shor's algorithm at scale would break RSA and elliptic-curve cryptography, the schemes securing most internet traffic, financial transactions, and communications today [1]. No such machine exists yet, but adversaries do not need one to begin profiting from this future: a "harvest now, decrypt later" strategy- intercepting and storing encrypted data today for decryption once quantum capability matures [13]- is already a documented concern for data with long confidentiality requirements, such as health and government records.

Post-quantum cryptography (PQC) addresses the cryptographic side of this problem [10], and adoption is underway in mainstream software- OpenJDK and OpenSSH have both shipped standardized ML-KEM support by default in recent releases [16]. What PQC standards do not address is sequencing: given finite migration resources, which data should move first, and how would an organization even know if harvesting is already happening against its systems? This paper treats that as a distinct, formalizable problem, building on AI techniques already established in cybersecurity for anomaly detection [6, 7]. Rather than stopping at a conceptual proposal, we test the resulting framework in simulation, find that our first design is flawed, and correct it- reporting that process rather than only the final result.

2. Literature Survey

2.1 Post-Quantum Cryptography: Technical Foundations

Shor's 1994 result motivated NIST's 2016 standardization competition [1, 10], which in July 2022 selected CRYSTALS-Kyber for key encapsulation [2] and CRYSTALS-Dilithium, FALCON, and SPHINCS+ for signatures. These rest on structured lattices and hash

functions rather than the factorization and discrete-logarithm problems underlying RSA and ECC [8], and were finalized as FIPS 203, 204, and 205 by August 2024 [10]. Benchmarks show Kyber and Dilithium performing competitively with, and in places faster than, classical RSA and ECDSA [2, 3], addressing early concerns about deployment cost.

Standardization does not guarantee implementation security, however. Wang, Xu, and Tian [5] extracted a full Kyber secret key from an ARM Cortex-M4 device using as few as 15 power traces via a side-channel attack- a reminder that mathematical resistance to quantum attack and resistance to physical implementation attack are separate properties.

2.2 Quantifying the Quantum Threat Timeline

This paper's risk model (Section 3) needs a probability distribution over CRQC arrival, and one is tracked systematically rather than guessed: the Global Risk Institute's Quantum Threat Timeline Report [12], run annually with evolutionQ since 2019, surveyed 26 experts in its 2025 edition and reported a 28–49% probability of a CRQC within 10 years and 51–70% within 15 - the highest 10-year estimate in the series. The shift is attributed partly to algorithmic progress: recent work has cut the estimated physical-qubit cost of breaking RSA-2048 by roughly 95%, from about 20 million to under one million qubits [11]. U.S. federal guidance (NIST IR 8547) correspondingly targets deprecating vulnerable algorithms after 2030 and disallowing them after 2035 [14], consistent with the NSA's CNSA 2.0 timeline for national security systems.

Section 4.6 derives calibration targets from these published GRI probability bands rather than assuming an arbitrary CRQC timeline.

2.3 The "Harvest Now, Decrypt Later" Threat in Practice

Palo Alto Networks [13] frames the arrival of a CRQC ("Q-Day") as a capability milestone, but notes the practical risk window opens well before it, since stolen ciphertext can simply wait. Separately, CrowdStrike's 2026 Global Threat Report found average eCrime breakout time- compromise to lateral movement- fell to 29 minutes in 2025 (from 48 in 2024), with the fastest observed at 27 seconds and one case reaching exfiltration in 4 minutes [15]. A slow-moving, decade-scale risk (quantum arrival) and a fast-moving, minute-scale risk (intrusion speed) together motivate the two-component architecture in Section 4: a QEW estimator updated infrequently, paired with an anomaly detector that must react quickly.

2.4 AI in Cybersecurity and the AI-PQC Intersection

Anomaly-based intrusion detection is well established as a complement to signature-based methods, which cannot catch attacks that don't match a known pattern [6, 7]. Autoencoder reconstruction error specifically — the basis for Section 4.3's detector — has a long track record in this literature, including recent work combining it with explainability layers so analysts can see why a flag was raised [7].

At the AI-PQC intersection specifically, Al-Mhadawi and Yas's review of 62 studies (2015–2025) found the field weighted heavily toward algorithm design and optimization (31.6%), cryptanalysis (26.3%), and side-channel defense (23.7%) [4], with comparatively little on operational detection or migration sequencing. Mothukuri and Parizi's 2026 review [11] makes a related observation: that quantum-computing and AI security research have largely developed as separate communities, and that deep learning can improve

side-channel key extraction even against protected implementations- a dual-use point Section 2.5 returns to. Their call for a defense-in-depth posture combining PQC, implementation hardening, and cryptographic agility is one this paper's framework contributes a piece of, specifically the prioritization and detection layer, without claiming to resolve the broader integration challenge.

2.5 AI as Threat: The Dual-Use Problem

The anomaly-detection technique proposed for defense in Section 4.3 belongs to the same technical family documented, by Al-Mhadawi and Yas's review [4], as already being used offensively against PQC: their review reports deep-learning power analysis demonstrated against Kyber and Dilithium implementations directly. Mothukuri and Parizi [11] describe such attacks as algorithm-agnostic at the implementation layer- meaning the detector proposed here should itself be designed with awareness that it could be evaded or have its training data poisoned by an adversary using comparable techniques [9]. *The reverse direction is also reported in the same review [4]: a CNN-based detector reaching over 98% accuracy (false-positive rate below 0.5%) specifically at detecting side-channel attacks against Kyber, rather than conducting them- direct precedent for the anomaly-detection approach this paper's Harvesting Detector (Section 4.3) is built on. We attribute both findings to the secondary review [4] rather than to the primary studies it cites: on manual inspection, the specific primary-source bibliographic entries for these two claims could not be independently verified (in one case, the cited page range belongs to an unrelated paper), so we do not reproduce them here as standalone references.*

2.6 Comparative Summary and Positioning

Table 1: Comparative summary of prior work and positioning relative to this paper.

Study / Direction	AI Role	Cryptographic Focus	Key Limitation
Al-Mhadawi & Yas [4]	Algorithm design, cryptanalysis, side-channel defense	PQC broadly	Detection/prioritization not treated as distinct problem
Mothukuri & Parizi [11]	Diagnoses AI/PQC split; discusses HNDL	PQC + AI cryptanalysis	Calls for integration but no formal prioritization
Almuhanna & Dardouri [6]; Sharma & Arora [7]	Anomaly detection via autoencoders, GNNs, LSTMs	Not PQC-specific	Not adapted to quantum-exposure prioritization
Wang et al. [5]	AI/signal processing as attack	Physical-layer PQC vulnerability	Demonstrates algorithm-agnostic implementation risk
Mosca & Piani (GRI) [12]	N/A (expert elicitation)	CRQC probability estimation	Provides empirical grounding for P(T Q)
This paper	Prioritization + detection, simulation-tested	Operational migration scheduling	Validated only on synthetic data (Section 4.6)

Taken together, the literature is more weighted toward cryptographic design and attack/defense than toward the operational question this paper asks: given finite resources, what gets migrated first, and how would an organization notice active harvesting at all? Recent reviews have begun pointing toward a need for closer integration between AI-security and quantum-security research [11]; this paper offers one concrete, testable mechanism toward that integration rather than claiming to settle it.

On prioritization versus blanket migration: a reasonable objection is that if PQC overhead keeps falling, organizations should simply migrate everything on a fixed schedule rather than build a scoring system. We think this understates two

separate problems: migration cost is not the only constraint (Section 3.2 lists three), and even where cost falls to near zero, organizational migration remains slow for reasons unrelated to compute cost- legacy systems, vendor dependencies, and certification cycles routinely stretch transitions over years regardless of the per-operation cost of the cryptography itself.

3. Problem Definition

3.1 The Harvest-Now-Decrypt-Later Threat Model

Let D denote a unit of encrypted data with sensitivity lifespan L_D - the number of years for which D must remain confidential (e.g., $L_D \approx 50$ for a medical record, $L_D \approx 2$ for

a stock trade confirmation). Let T_Q denote the unknown time until a cryptographically relevant quantum computer (CRQC) becomes operational, estimable from expert-elicitation series such as the GRI timeline [12]. An adversary intercepts D at time t_0 and can decrypt it at any $t \geq T_Q$, so D is compromised under quantum risk if:

$$T_Q \leq t_0 + L_D$$

Interpretation. Before introducing the risk quantity below, we fix its meaning precisely, since "exposure" could otherwise be read several ways. Under uncertainty in T_Q , we define the Quantum Exposure Window (QEW) as the expected duration of overlap between an asset's remaining confidentiality lifetime and the period in which a CRQC may exist- an expectation taken over the full distribution of T_Q , not a single assumed arrival time. QEW is therefore an expected number of years of overlap, not a probability and not a raw lifespan. As an intuitive point-estimate approximation, QEW can first be written as:

$$\text{QEW}(D) = \max(0, t_0 + L_D - E[T_Q]) \quad (3.1)$$

Because this point-estimate form, Eq. (3.1), collapses the full uncertainty in T_Q down to a single value $E[T_Q]$, it is only an approximation to the definition above; Section 4.2 replaces it with Eq. (4.2), the distributional formulation actually used in the simulation (Section 4.6), integrating over the survival probability of T_Q rather than assuming a single arrival time. Here, QEW near zero means the data's confidentiality naturally expires before quantum risk materializes; a large QEW means years of required confidentiality overlap with plausible quantum capability, and should be prioritized accordingly.

3.2 Why Blanket PQC Migration Is Insufficient

Migrating all encrypted data to PQC immediately would eliminate QEW, but three constraints make this impractical:

- 1) **Computational overhead-** PQC algorithms generally require larger keys and more computation than classical schemes, which matters disproportionately for constrained environments [4].
- 2) **Migration cost scales with volume-** organizations often hold petabytes of historical encrypted data; simultaneous re-encryption is rarely feasible.
- 3) **Detection blindness-** migrating forward does not reveal whether harvesting of historical data has already occurred [11, 13].

Problem Statement: Given a corpus $\{D_1, \dots, D_n\}$ with varying sensitivity lifespans and observable network telemetry, estimate $\text{QEW}(D_i)$ for each asset and detect behavioral signatures consistent with active harvesting [6, 7], producing a prioritized migration schedule rather than uniform blanket migration.

3.3 Detection as a Distinct Sub-Problem

Let $\mathbf{x}_t \in \mathbb{R}^k$ be a feature vector of network/system behavior at time t . A harvesting event should shift $\mathbf{x}_t$'s distribution relative to a learned baseline $P(\mathbf{x})$ of normal behavior [6, 7]:

$$\text{Anomaly Score}(\mathbf{x}_t) = -\log P(\mathbf{x}_t | \theta)$$

where θ are the parameters of a model trained on historical benign behavior- Section 4 operationalizes this as an autoencoder.

4. Methodology / Approach

4.1 System Overview

Two components operate over an organization's data corpus and telemetry: a QEW Estimator scoring each asset from metadata, and a Harvesting Detector monitoring telemetry $\mathbf{x}_t$ for exfiltration signatures. Both feed a Prioritized Migration Scheduler that ranks assets for re-encryption.

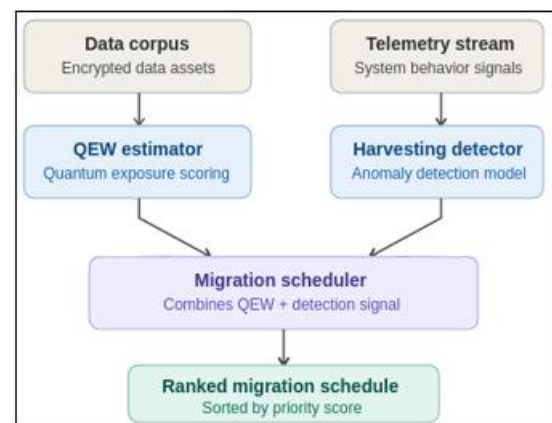


Figure 1: Architecture of the Prioritized PQC Migration Scheduling Framework

4.2 QEW Estimator

Given lifespans L_{D_i} and a CRQC arrival distribution $P(T_Q)$ informed by expert forecasts [12], the point-estimate form in Section 3.1 assumes a single expected arrival $E[T_Q]$. Forecast uncertainty is substantial enough that we replace this with an integral over the survival probability $P(T_Q > \tau)$:

$$\text{QEW}(D_i) = \max\left(0, L_{D_i} - \int_0^{L_{D_i}} P(T_Q > \tau) d\tau\right) \quad (4.2)$$

This accounts for the probability the CRQC has not yet arrived by time τ , rather than collapsing a genuinely uncertain forecast into one number.

4.3 Harvesting Detector

Detection is formalized as unsupervised anomaly detection: an autoencoder f_θ reconstructs benign traffic, with reconstruction error as the anomaly signal. Under a Gaussian error assumption, minimizing this term is equivalent, up to a constant, to maximizing the log-likelihood in Section 3.3 [6, 7]:

$$\mathcal{L}_{\text{recon}}(\mathbf{x}_t) = \|\mathbf{x}_t - f_\theta(\mathbf{x}_t)\|_2^2 \quad (4.3)$$

An event is flagged when $\mathcal{L}_{\text{recon}}(\mathbf{x}_t)$ exceeds τ_{anomaly} , calibrated on held-out data for a target false-positive rate. Section 4.6 reports the actual precision, recall, and FPR achieved.

Scope assumption: asset-to-telemetry attribution

Steps 7–8 of Algorithm 1 (Section 4.4) flag "associated data assets" once an anomaly is detected - this attribution is assumed, not solved, by the present framework. A real deployment would need to link a flagged telemetry event to specific data assets via flow logs, object/database access logs, destination metadata, or DLP labels, none of which this paper models. We treat the existence of such an attribution layer as a precondition for deployment and flag building it as separate, non-trivial engineering work, distinct from the scoring and detection logic this paper formalizes.

4.4 Prioritized Migration Scheduler- Algorithm (Revised)

This section reflects a correction made after simulation testing (Section 4.6). An earlier version combined QEW and the harvesting signal multiplicatively:

$$\text{priority}(D_i) = \text{QEW}(D_i) \times \text{HARVEST_BOOST}$$

Testing this in simulation (Section 4.6) exposed a structural flaw: multiplying a near-zero QEW by any constant boost stays near zero, so an actively harvested asset with a short lifespan remains low-priority regardless of boost size- even a $50\times$ boost caught under 10% of such cases at a realistic migration budget. We replace the rule with an additive combination over the percentile rank of QEW:

$$\text{priority}(D_i) = \text{PctRank}(\text{QEW}(D_i)) + \mathbb{I}[D_i \in \text{Harvested}] \quad (4.4)$$

where $\text{PctRank}(\text{QEW}(D_i)) \in [0, 1]$ is the fraction of the corpus with strictly lower QEW, and the indicator adds a full unit when D_i is flagged.

Algorithm 1: Prioritized PQC Migration Scheduling (Revised)

```

Input: Data corpus {D_1, ..., D_n}
       Telemetry stream {x_t}
       CRQC arrival distribution P(T_Q)
       Anomaly threshold  $\tau_{\text{anomaly}}$ 
Output: Ranked migration schedule S

1: for each D_i in corpus:
2:   compute QEW(D_i) using Eq. (4.2)
3: end for
4: compute PctRank(QEW(D_i)) for all D_i // rank over the corpus
5: for each x_t in telemetry stream:
6:   compute L_recon(x_t) using Eq. (4.3)
7:   if L_recon(x_t) >  $\tau_{\text{anomaly}}$ :
8:     flag associated assets as ACTIVELY_HARVESTED // see 4.3 scope note
9:   end if
10: end for
11: for each D_i in corpus:
12:   if D_i is ACTIVELY_HARVESTED:
13:     priority(D_i)  $\leftarrow$  PctRank(QEW(D_i)) + 1
14:   else:
15:     priority(D_i)  $\leftarrow$  PctRank(QEW(D_i))
16:   end if
17: end for
18: S  $\leftarrow$  sort corpus by priority(D_i) descending
19: return S

```

Proposition 1. Under the additive percentile-rank priority rule, every flagged asset ranks strictly above every non-flagged asset, regardless of its QEW.

Proof: For any non-flagged D_j , $\text{priority}(D_j) = \text{PctRank}(\text{QEW}(D_j)) < 1$. For any flagged D_i , $\text{priority}(D_i) = \text{PctRank}(\text{QEW}(D_i)) + 1 \geq 1$. Hence $\text{priority}(D_i) \geq 1 > \text{priority}(D_j)$ for every non-flagged D_j .

Caveat: this guarantee is detector-dominant by design.

Proposition 1 is a clean mathematical guarantee, but it has a direct operational consequence worth stating plainly: if the detector produces false positives, every false positive is guaranteed to outrank every true high-sensitivity asset that was not flagged, by construction. The binary indicator makes no distinction between a confident detection and a borderline one. In a high-false-positive environment, this could let noise consume the entire migration budget. A natural refinement replaces the binary indicator with a continuous, confidence-weighted term $p_i \in [0, 1]$ - the detector's estimated harvesting probability for D_i - rather than a 0/1 flag:

$$\text{priority}(D_i) = \text{PctRank}(\text{QEW}(D_i)) + p_i, p_i \in [0, 1]$$

This preserves the spirit of Proposition 1 for high-confidence detections while letting low-confidence flags compete with QEW on a graded basis rather than automatically dominating it. We did not implement this variant in the Section 4.6 simulation and flag it as a direct extension for future work (Section 7).

4.5 Evaluation Plan

Section 4.6 benchmarks the detector's precision/recall against a calibrated false-positive target and compares the Scheduler's ranking quality against a naive "migrate by lifespan" baseline, following conventions in the intrusion-detection benchmarking literature [6, 7].

4.6 Simulation Study

This section reports results from a synthetic simulation, not a real-world deployment, intended to stress-test the framework's internal logic under controlled, known ground truth.

Setup: We generated 500 synthetic data assets across five categories (financial transactions, health records, government communications, intellectual property, general logs), each with a category-appropriate sensitivity lifespan. $P(T_Q)$ was fitted as a Weibull distribution- a standard choice for time-to-event modeling because its shape parameter k flexibly captures accelerating, constant, or decelerating hazard rates without assuming a specific mechanism. Since GRI reports a range rather than a point estimate at each horizon, we selected target probabilities near the midpoint of each reported range (38.5% at 10 years, within the 28–49% band; 60.5% at 15 years, within the 51–70% band [12]) and solved numerically via root-finding for the (k, λ) pair reproducing both targets, giving $k = 1.60$, $\lambda = 15.71$. The resulting CDF gives approximately 38.5% probability by 10 years and 60.5% by

15 years- midpoint-consistent with, not identical to, the full GRI range. QEW was computed via Eq. (4.2) for every asset. Synthetic telemetry used 12 features per asset, intended as stand-ins for coarse, commonly logged network signals- for example, egress volume, number of distinct destination endpoints, access frequency to archived/encrypted objects, and time-of-day irregularity- generated as benign multivariate-normal data. 15% of assets were randomly marked ground-truth "actively harvested," independent of lifespan category, and given an injected shift in 4 of the 12 features to simulate a bulk-egress signature. We emphasize that this is a deliberately clean separation between benign and harvesting classes; Section 5.3 discusses why this likely overstates achievable real-world detector performance.

Detector performance: An autoencoder (8-3-8 architecture) trained on held-out benign telemetry, with τ_{anomaly} calibrated for a 1% false-positive target, achieved precision = 0.951, recall = 0.784, $F1 = 0.859$, and an observed FPR of 0.7% on the full corpus- close to target, confirming the calibration procedure behaves as intended under these (favorable) synthetic conditions.

Scheduler ranking quality- the central result. We isolated the 41 synthetic assets that were both actively harvested and below-median QEW- short-lifespan data under active attack, the case a lifespan-only policy structurally cannot catch. At a top-20%-of-corpus migration budget, the naive baseline and the original multiplicative scheduler (boost = 2.0) both caught 0% of these; a 50 $\times$ boost caught 9.8%. The revised additive scheduler caught 75.6% at the identical budget.

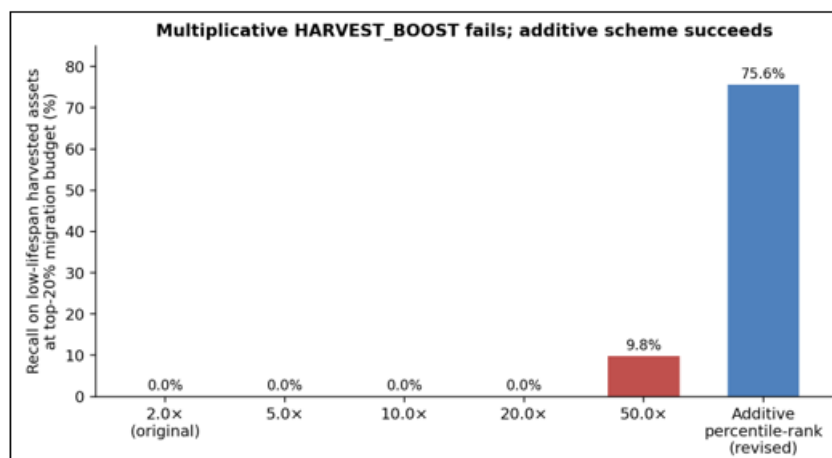


Figure 2: Recall on short-lifespan, actively-harvested assets at a top-20% migration budget. The multiplicative design fails regardless of boost magnitude; the additive scheme (Proposition 1) succeeds.

This comparison- not detector accuracy- is the paper's central empirical contribution: it shows concretely why the additive rule is a necessary correction rather than a stylistic preference, and the scheduler comparison is more informative here than the detector metrics above, since the detector's strong numbers reflect a favorable synthetic separation rather than a claim about real-network difficulty.

5. Results and Discussion

5.1 What the Simulation Shows and Does Not Show

The simulation confirms two things with actual numbers rather than argument alone: the detector operates near its calibration target (0.7% observed vs. 1% target FPR), and the additive scheduler closes a real gap the multiplicative design left open. It does not show that the framework works on real telemetry, that the 12 synthetic features resemble real log data closely enough, or that detection and QEW combine this cleanly once the asset-attribution problem (Section 4.3) is actually solved rather than assumed.

With GRI data grounding $P(T_Q)$ [12], the framework converts migration from a calendar-driven decision into a risk-driven one layered on top of regulatory deadlines (Section 2.2)- using the mandated deadline as an upper bound and the QEW/harvesting signal, now provably well-combined under Proposition 1's stated assumptions, to sequence effort within it.

5.2 Position Relative to Prior Work

Relative to the AI-for-PQC-design literature [4], this work is complementary: it assumes PQC algorithms already exist and addresses sequencing, not design. Relative to Mothukuri and Parizi's observation that AI-security and quantum-security research have developed separately [11], this paper contributes one tested mechanism at that junction- a bounded contribution, since Section 4.6's validation is synthetic and the dual-use risk identified in Section 2.5 remains unaddressed by the simulation, which assumed a non-adversarial telemetry generator.

A further limitation: GRI's own reports note their estimates lag fast-moving theoretical progress [12], a property the simulation's single-snapshot Weibull fit does not capture. A deployed system would need $P(T_Q)$ refreshed on each new report rather than fitted once.

5.3 Practical Limitations

- 1) **Synthetic validation only.** Section 4.6 uses simulated telemetry with a clean, injected signal. Real telemetry is noisier, less separable, and not independent of lifespan category in the way the simulation assumes.
- 2) **Asset-to-telemetry attribution is assumed, not solved.** As noted in Section 4.3, mapping a flagged anomaly to a specific data asset requires infrastructure (flow logs, DLP labels, access logs) this paper does not model.
- 3) **Ground-truth scarcity.** Real deployments need labeled exfiltration examples [6, 7], which are rare and organization-specific.
- 4) **Adversarial vulnerability untested.** Section 2.5's dual-use findings [4, 9, 11] suggest the detector could itself be targeted; the simulation did not test this.
- 5) **Detector-dominant ranking.** As discussed after Proposition 1, false positives automatically outrank true high-sensitivity data under the binary indicator; the confidence-weighted extension (Eq. in Section 4.4) is untested.
- 6) **Forecast staleness and untuned architecture.** $P(T_Q)$ is a static fit rather than a refreshed one, and the autoencoder used a single, untuned configuration.

6. Conclusion

This paper formalized a prioritization and detection problem for post-quantum migration- which data to protect first, and how to notice active harvesting- building on a literature that has largely treated quantum-computing and AI-security research as separate concerns [11]. We proposed a Quantum Exposure Window metric grounded in published expert forecasts [12], paired with a behavioral anomaly detector.

The simulation in Section 4.6 served not only as validation but as a design test. It exposed a genuine limitation of the originally proposed multiplicative scheduler- a failure to prioritize short-lifespan data under active attack- and motivated the revised additive ranking rule, which we proved closes that gap under stated assumptions and confirmed empirically (0% to 75.6% recovery on the affected cases). That same testing process also surfaced the rule's detector-dominant character, an assumption about asset-telemetry attribution this paper does not resolve, and the synthetic nature of the validation itself- limitations we have tried to state plainly rather than minimize.

7. Future Scope

- 1) **Real-world validation** of both the detector and scheduler against production or realistic telemetry (e.g., CICIDS-derived, relabeled for exfiltration), replacing the synthetic generator in Section 4.6.
- 2) **Building the asset-to-telemetry attribution layer** flagged in Section 4.3 as assumed rather than solved — likely requiring integration with flow logs, DLP tooling, or database access auditing.
- 3) **Implementing and testing the confidence-weighted priority rule** (Section 4.4) to address the detector-dominant ranking behavior identified after Proposition 1.
- 4) **Adversarial robustness testing** [9], given documented offensive use of structurally similar AI techniques against PQC implementations [4, 5].
- 5) **Dynamic $P(T_Q)$ refresh** tied to each new GRI report [12], replacing the static Weibull snapshot used in Section 4.6.
- 6) **Explainable AI integration**, responding to the gap named as most critical in the field's largest existing systematic review [4].

References

- [1] Shor PW. Algorithms for quantum computation: discrete logarithms and factoring. Proceedings of the 35th Annual Symposium on Foundations of Computer Science; 1994. p. 124–134.
- [2] Bos J, Ducas L, Kiltz E, Lepoint T, Lyubashevsky V, Schanck JM, et al. CRYSTALS – Kyber: a CCA-secure module-lattice-based KEM. Proceedings of IEEE EuroS&P; 2018. p. 353–367.
- [3] Ducas L, Kiltz E, Lepoint T, Lyubashevsky V, Schwabe P, Seiler G, et al. CRYSTALS-Dilithium: a lattice-based digital signature scheme. IACR Trans. on Cryptographic Hardware and Embedded Systems. 2018;2018(1):238–268.
- [4] Al-Mhadawi MM, Yas QM. Quantum-safe artificial intelligence: a systematic review of post-quantum cryptography applications. Int. J. of Current Science Research and Review. 2026;9(6):3002–3015.
- [5] Wang K, Xu D, Tian J. An improved two-step attack on lattice-based cryptography: a case study of Kyber. arXiv preprint (not peer-reviewed). 2024. Available from: <https://arxiv.org/abs/2407.06942>
- [6] Almuhananna R, Dardouri S. A deep learning/machine learning approach for anomaly based network intrusion detection. Frontiers in Artificial Intelligence. 2025; 8: 1625891.

- [7] Sharma N, Arora B. Machine learning and deep learning models for anomaly intrusion detection in networks: a systematic review. *SN Computer Science*. 2025; 6: 832.
- [8] Peikert C. A decade of lattice cryptography. *Foundations and Trends in Theoretical Computer Science*. 2016;10(4):283–424.
- [9] Biggio B, Roli F. Wild patterns: ten years after the rise of adversarial machine learning. *Pattern Recognition*. 2018; 84: 317–331.
- [10] NIST. Post-Quantum Cryptography Program [Internet]. 2024. Available from: <https://csrc.nist.gov/Projects/post-quantum-cryptography>
- [11] Mothukuri V, Parizi RM. Securing cryptography in the age of quantum computing and AI: threats, implementations, and strategic response. *arXiv preprint (not peer-reviewed)*. 2026. Available from: <https://arxiv.org/abs/2603.06969>
- [12] Mosca M, Piani M. Quantum Threat Timeline Report 2025. *Global Risk Institute / evolutionQ*; 2026 Mar. Available from: <https://globalriskinstitute.org/publication/quantum-threat-timeline-report-2025b/>
- [13] Palo Alto Networks. What is Q-Day? Quantum computing and cyber risk [Internet]. 2026.
- [14] National Institute of Standards and Technology. NIST IR 8547 (Initial Public Draft): Transition to Post-Quantum Cryptography Standards. 2024 Nov. Available from: <https://csrc.nist.gov/pubs/ir/8547/ipd>
- [15] CrowdStrike. 2026 Global Threat Report [Internet]. 2026 Feb. Available from: <https://www.crowdstrike.com/global-threat-report/>
- [16] OpenJDK. JEP 496: Quantum-Resistant Module-Lattice-Based Key Encapsulation Mechanism [Internet]. 2024. Available from: <https://openjdk.org/jeps/496> ; OpenSSH Project. Release Notes, OpenSSH 10.0 [Internet]. 2025 Apr. Available from: <https://www.openssh.com/txt/release-10.0>