Journal of Computer Science and Cybernetics, V.41, N.x (2025), 1–
DOI no. 10.15625/1813-9663/23150

OPTIMIZING PACKET SIZE IN POST-QUANTUM NB-IOT SYSTEMS: SIGNATURE AGGREGATION AND MERKLE TREE PRUNING APPROACHES

THI BAC DO, KHANH LINH DINH*

Thai Nguyen University of Information and Communication Technology, Z115 Street, Quyet Thang Ward, Thai Nguyen Province, Viet Nam

Abstract. With the increasing adoption of post-quantum cryptographic schemes such as SPHINCS+ in Narrowband IoT (NB-IoT) systems, a major challenge arises from large packet size overhead, often reaching approximate 34KB per transmission due to stateless hash-based signature metadata. This paper investigates practical techniques for reducing this overhead while maintaining quantum resilience, including signature aggregation across multiple devices, Merkle tree pruning within SPHINCS+, and selective signing policies based on session frequency. We implement these strategies on an ESP32-WROOM-32 microcontroller interfaced with a SIM7080 NB-IoT module, demonstrating that packet size can be reduced by up to 60% through intelligent session management and optimized scheduling. While our system integrates Kyber for key exchange and ChaCha20-Poly1305 for encryption to form a complete hybrid post-quantum secure channel, the optimization focus remains exclusively on mitigating the dominant source of packet expansion: SPHINCS+ signatures. These components, though essential for end-to-end security, contribute negligibly to payload inflation compared to the signature metadata. By narrowing the scope to this critical bottleneck, our work provides actionable insights for system architects aiming to deploy post-quantum authentication within the strict bandwidth constraints typical of LPWAN environments.

Keywords. Post-quantum digital signatures, signature aggregation, Merkle tree pruning, packet size optimization, NB-IoT security, ChaCha20-Poly1305, offloading verification, selective signing policy, lightweight IoT security, hybrid authentication mechanisms.

1. INTRODUCTION

As post-quantum cryptographic algorithms become increasingly standardized, their deployment on constrained Internet of Things (IoT) platforms presents new challenges beyond computational feasibility, one of which is packet size expansion due to large digital signatures. To address this issue, some research focuses on improving algorithms or algebraic foundations to reduce the size of signatures while still ensuring security [1,2]. Other research focuses on optimizing the cost of algorithms [3–5]. This paper focuses specifically on mitigating the significant overhead introduced by SPHINCS+, a hash-based digital signature scheme selected by NIST for its long-term security guarantees.

*Corresponding author.
E-mail addresses: dtbac@ictu.edu.vn(T.B. Do); dklinh@ictu.edu.vn(K.L. Dinh).

In narrowband IoT (NB-IoT) networks, where typical payloads are limited to 1 - 2 KB per transmission, the integration of SPHINCS+ introduces metadata that can reach up to ∼34KB per packet, far exceeding acceptable limits [6]. While this level of security is essential for future-proofing data confidentiality, it also poses practical limitations in real-world deployment scenarios.

To address this challenge, we propose a set of optimization strategies including signature aggregation, Merkle tree pruning, and selective invocation of PQC operations, all aimed at reducing packet overhead while maintaining resistance against both classical and quantum adversaries. These techniques are implemented and evaluated on an ESP32-WROOM-32 microcontroller interfaced with a SIM7080 NB-IoT module, measuring execution time, memory usage, and power draw under realistic network conditions.

Unlike prior work that emphasized full integration of PQC layers, our study narrows its scope to one of the most pressing limitations introduced by SPHINCS+: packet-level inefficiency. We build upon earlier research by shifting focus toward actionable mitigation strategies that improve deployability without compromising long-term security guarantees.

The rest of this paper is organized as follows: Section 2 provides background on SPHINCS+ and its impact on NB-IoT packet size; Section 3 presents related work on signature reduction techniques; Section 4 details our proposed optimizations; Section 5 evaluates them experimentally; Section 6 discusses implications for industrial deployment; and Section 7 concludes with recommendations for future research directions.

2. BACKGROUND AND MOTIVATION

The widespread adoption of Narrowband IoT (NB-IoT) in smart city infrastructure has made it a critical platform for long-term data transmission in remote sensing, smart metering, and environmental monitoring applications [7]. However, as quantum computing becomes increasingly viable, traditional cryptographic schemes such as RSA and ECC are expected to become obsolete within the next few decades due to their vulnerability to Shor's algorithm [8]. This necessitates a transition toward post-quantum cryptography (PQC), particularly in systems where firmware updates are rare or impossible.

Among the standardized algorithms selected by NIST, SPHINCS+, a stateless hash-based digital signature scheme, is one of the most promising candidates for securing constrained IoT devices without relying on complex state management or hardware acceleration [6]. Unlike lattice-based schemes such as Dilithium or Falcon, which require careful tracking of signing operations and are vulnerable to side-channel leakage [9], SPHINCS+ offers a stateless design that simplifies deployment on small-footprint platforms like ESP32 or STM32WB.

However, despite its robustness against quantum adversaries, SPHINCS+ introduces a significant drawback: packet size overhead. A typical SPHINCS+ signature can reach up to ∼34KB, far exceeding the payload limits of NB-IoT networks, which typically operate under strict constraints of 1 - 2KB per transmission [10]. This makes SPHINCS+ unsuitable for direct use in many LPWAN environments unless mitigation strategies are applied.

This paper focuses specifically on addressing this issue, the challenge of integrating SPHINCS+ into NB-IoT while maintaining compatibility with bandwidth-limited communication. We explore practical optimization techniques, including signature aggregation, Merkle tree pruning, and selective signing policies, all aimed at reducing the impact of large signatures without compromising quantum resilience.

2.1. Why SPHINCS+ causes large packet overhead

SPHINCS+ is based on a hierarchical Merkle tree structure, which generates multiple layers of authentication paths during signature creation [10]. The default configuration, SPHINCS+-SHA256-128f, results in signatures of approximately 34 KB, which is orders of magnitude larger than classical schemes like ECDSA (∼64-128 bytes). This expansion stems from two main factors:

- Stateless design: Unlike Dilithium or Falcon, which reuse internal states across signing operations, SPHINCS+ must include full path information in each signature.
- Security level requirements: To resist quantum attacks, SPHINCS+ relies on deep trees and large hash outputs, increasing metadata significantly.

As shown in recent studies, this level of overhead severely impacts NB-IoT network efficiency, especially in deployments where frequent transmissions are not feasible due to limited bandwidth and power budget [7].

Unlike earlier work that emphasized computational feasibility or memory footprint, our focus lies specifically on packet-level efficiency, an aspect largely overlooked in previous literature. We build upon findings by Bürstinghaus-Steinbach et al. [11], who evaluated SPHINCS+ performance on microcontroller platforms, but extend their analysis to real-world deployment challenges in NB-IoT environments.

2.2. Impact on NB-IoT deployments

Narrowband IoT (NB-IoT) is widely used in urban sensing applications due to its low-power operation and broad coverage [12]. It typically transmits small amounts of data infrequently, once every few hours or days, making it ideal for battery-powered devices with multi-year lifecycles.

However, the integration of PQC increases packet size beyond acceptable limits for standard NB-IoT modules. As demonstrated by Bernstein et al. [13], deploying SPHINCS+ without any optimization leads to transmission inefficiencies, increased latency, and higher energy consumption, all of which hinder real-world deployability.

To mitigate these issues, we propose several techniques:

- Signature aggregation allows multiple devices to share a single signature block.
- Tree pruning, reducing the depth of Merkle structures while preserving security guarantees.
- Selective signing policy, invoking SPHINCS+ only when necessary.
- Offloading verification, shifting heavy computation to gateway-level processing units.

These strategies aim to reduce packet overhead while ensuring that the system remains secure against both classical and quantum adversaries.

Unlike lattice-based schemes such as Dilithium or Falcon, which require complex state management and are vulnerable to timing attacks [14], our study demonstrates how stateless hash-based signatures can be optimized for real-world NB-IoT deployments, offering a practical balance between quantum resistance and network efficiency.

3. RELATED WORK

The deployment of post-quantum cryptographic schemes on constrained Internet of Things (IoT) platforms has gained increasing attention as quantum computing becomes more viable [6]. However, most existing studies focus on evaluating individual PQC algorithms or proposing full-stack secure communication frameworks, often overlooking one of the most critical limitations introduced by hash-based digital signatures like SPHINCS+: packet size expansion.

Several researchers have explored the feasibility of implementing PQC on microcontroller platforms such as ARM Cortex-M4, ESP32, or STM32WB55RG. For example, Bürstinghaus-Steinbach et al. [11] evaluated SPHINCS+ on Cortex-M4 microcontrollers and showed that while it is a feasible option for long-term secure communication systems, its signature size (∼34KB) significantly exceeds the typical payload limit (∼1-2KB) of NB-IoT networks [7].

Valdivieso et al. [15] further analyzed the implementation challenges of SPHINCS+ on embedded platforms, emphasizing the need for manual memory management and constant-time execution to reduce timing vulnerabilities. Their work highlights that while SPHINCS+ offers strong security guarantees, its integration into LPWAN technologies must be accompanied by techniques to reduce metadata overhead.

Zaverucha et al. [16] proposed signature aggregation, allowing multiple devices to share a single digital signature in certain deployment scenarios. This method reduces total network load by combining authentication data from multiple nodes, making it particularly useful in urban sensing applications where groups of sensors transmit simultaneously.

Another promising approach is Merkle tree pruning, which involves limiting the depth of the Merkle structure used in SPHINCS+ to reduce signature length. As shown in prior studies, this strategy can lower packet size by up to 40–50% while still maintaining resistance against forgery attacks.

In contrast, Rana et al. [17] presented a survey on lightweight cryptographic techniques suitable for constrained IoT environments, highlighting the importance of balancing efficiency and security. While their work focuses primarily on classical schemes, it supports the idea that system-level optimizations, rather than algorithmic novelty alone, should guide practical deployment decisions.

Bürstinghaus-Steinbach et al. [11] discussed practical deployment strategies for post-quantum signatures on embedded systems, advocating for simplified signing processes and fallback mechanisms during transition periods. We build upon this concept by introducing concrete technical solutions specifically targeting packet-level efficiency.

Unlike earlier efforts that emphasized theoretical proofs or isolated performance benchmarks, our study centers on actionable methods for reducing packet overhead caused by large digital signatures, ensuring compatibility with real-world NB-IoT constraints.

While Gupta et al. [18] tested Kyber and NewHope on small sensor platforms and concluded that PQC integration is feasible with effective software optimization, they did not address the issue of signature size, a gap we aim to fill through targeted mitigation strategies.

Similarly, Greconici et al. [19] assessed the performance of Kyber and Dilithium on Cortex-M4 microcontrollers but focused only on computational cost, leaving packet size considerations unexplored.

This paper builds on previous findings and introduces new contributions centered around packet-level efficiency, providing firmware developers and system architects with practical insights for deploying SPHINCS+ in resource-constrained NB-IoT environments.

As standardization efforts by NIST and ETSI accelerate, transitioning toward post-quantum infrastructure is no longer a distant requirement; it is something we must act upon today, especially for systems with multi-year deployment cycles.

4. PROPOSED OPTIMIZATION TECHNIQUES

The deployment of post-quantum cryptographic schemes on Narrowband IoT (NB-IoT) platforms introduces significant challenges, particularly in terms of packet size expansion due to large digital signatures. One of the most notable limitations arising from SPHINCS+ is a stateless hash-based digital signature scheme selected by NIST for its resistance to both classical and quantum adversaries [8].

While SPHINCS+ offers robust security guarantees, it generates metadata that can reach up to ~34KB per transmission, far exceeding the typical payload limit (~1 - 2KB) of NB-IoT networks [17]. To address this issue while maintaining post-quantum resilience, we propose a set of practical optimization techniques specifically designed to reduce the impact of PQC components on network performance. These include:

- Signature aggregation across multiple devices.
- Merkle tree pruning within SPHINCS+.
- Selective invocation of PQC operations based on session frequency.
- Offloading signature verification to the gateway or cloud services.

These strategies are implemented and evaluated on an ESP32-WROOM-32 microcontroller interfaced with a SIM7080 NB-IoT module, ensuring compatibility with existing LP-WAN protocols while minimizing unnecessary cryptographic overhead.

4.1. Signature aggregation across multiple devices

One effective way to reduce packet size expansion is to apply signature aggregation, allowing multiple NB-IoT nodes to share a single digital signature during synchronized transmissions. This technique is especially useful in smart city applications where groups of sensors send environmental readings or metering data simultaneously [16].

In our framework, we implement a lightweight aggregation mechanism based on the hierarchical structure of SPHINCS+, enabling multiple signatures to be combined into a single block without compromising authenticity or forward secrecy. As shown by Zaverucha et al. [16], such approaches significantly reduce total network load, potentially lowering aggregate overhead by up to 6%, making them ideal for bandwidth-limited environments.

This method aligns with findings by Bürstinghaus-Steinbach et al. [11], who advocated for simplified signing processes and fallback mechanisms in embedded systems. We build upon this idea by introducing concrete technical solutions specifically targeting packet-level efficiency.

Unlike earlier work that focused solely on individual device authentication, our study demonstrates how aggregation can be applied in real-world NB-IoT deployments, reducing packet overhead while preserving quantum resilience.

4.2. Merkle tree pruning within SPHINCS+

A major reason for the large packet overhead in SPHINCS+ is the depth of the Merkle trees used during signature generation [20]. Each node in the tree contributes additional authentication paths, increasing the final signature length.

To mitigate this, we apply tree pruning, a method that reduces the height of the Merkle structure without compromising security beyond acceptable limits. Instead of using the default configuration of SPHINCS+-SHA256-128f, which results in a 34 KB signature, we deploy the compact variant SPHINCS+-SHA256-128s, reducing signature size to approximately 18 KB per packet.

While this adjustment lowers the security category from NIST's Category 4 to Category 3, it remains resistant to all known attacks for the next 10 - 15 years, sufficient for many industrial IoT deployments [20].

Zhou et al. [20] have explored similar optimizations on embedded platforms and concluded that stateless hash-based signatures remain viable even under reduced Merkle tree depths, provided they are deployed with appropriate session management policies.

Our implementation builds on these insights by demonstrating how tree pruning can be applied directly on NB-IoT modules like ESP32, offering a practical solution for organizations seeking to adopt PQC without overhauling their entire communication stack.

4.3. Selective signing policy based on session frequency

Rather than executing full digital signatures on every transmission, we introduce a selective signing policy, where SPHINCS+ is invoked only when necessary, such as during initial authentication or when critical data is sent.

For non-critical or periodic sensor readings, previously verified session keys are reused, avoiding redundant computation and reducing overall packet overhead. This approach builds upon work by Abbasi et al. [21], who emphasized the importance of session reuse in resource-constrained environments.

We further enhance this method by incorporating nonce-based freshness checks, ensuring that even when signatures are reused, messages remain unique and verifiable. This helps prevent replay attacks while still reducing the frequency of expensive cryptographic operations.

By applying selective signing, the system maintains strong resistance to forgery and impersonation, yet significantly improves energy efficiency and scalability, particularly in environments where firmware updates are rare and power conservation is critical [21].

4.4. Offloading signature verification to the gateway or cloud services

Given the limited processing capabilities of NB-IoT nodes, offloading certain cryptographic operations to more powerful edge or cloud services is a promising solution. In particular, we explore shifting signature verification from the client side to the gateway or backend server, which typically has greater computational capacity and hardware acceleration support.

Under this model:

- The NB-IoT client sends only the public key and message hash.

- The gateway performs the full signature verification using optimized libraries like liboqs or wolfSSL.

- Clients benefit from reduced processing load and lower power draw during authentication phases.

This method ensures compatibility with legacy NB-IoT modules while enabling gradual migration toward full PQC deployment. It also allows organizations to update backend infrastructure before rolling out firmware changes to remote devices, a realistic approach given the long deployment cycles of NB-IoT networks [14].

Bürstinghaus-Steinbach et al. [11] have discussed similar offloading mechanisms in their work on practical deployment of PQC in embedded systems, emphasizing the importance of distributing cryptographic workload based on platform capabilities.

Table 1: Comparison of optimization techniques

Technique	Packet size reduction	Quantum resilience	Hardware dependency	Energy efficiency
Original SPHINCS+	No	Strong	High	Medium
Signature aggregation	60%	Strong	Low	High
Merkle tree pruning	40–50%	Moderate	Low	High
Selective signing	30–40%	Strong	Low	Very High
Offloading verification	90% (on client side)	Strong (on gateway)	Low	Very High

As summarized in Table 1, each optimization technique presents different trade-offs between packet reduction, quantum resistance, hardware dependency, and energy efficiency. Our study does not advocate for any single method but instead proposes a flexible model where these strategies can be applied individually or in combination, depending on deployment requirements.

5. IMPLEMENTATION AND EVALUATION

To systematically evaluate the effectiveness of the proposed optimization techniques in mitigating packet size overhead induced by post-quantum cryptographic digital signatures, a series of controlled experiments was conducted utilizing an ESP32-WROOM-32 microcontroller coupled with a SIM7080 NB-IoT module. Unlike comprehensive end-to-end evaluations performed in hybrid PQC frameworks, as exemplified by prior work [14], our investigation explicitly concentrates on analyzing and addressing the substantial overhead introduced by SPHINCS+ metadata through targeted engineering methodologies. We implemented and tested four optimization approaches.

Four distinct optimization strategies were implemented and thoroughly assessed:

- Signature aggregation.

- Merkle tree pruning.

- Selective signing policy.
- Verification offloading to gateway.

Each approach was scrutinized under realistic operating conditions characteristic of typical NB-IoT networks, which include infrequent firmware updates, stringent power-saving requirements, and constrained bandwidth.

5.1. Experimental setup and measurement methodology

Table 2 provides a comprehensive overview of the hardware and software configuration used for our experimental evaluation.

Table 2: Experimental Setup and Hardware Configuration

Component	Description
Microcontroller	ESP32-WROOM-32, Wi-Fi and Bluetooth-enabled chip with ∼520 KB SRAM and 4 MB Flash
NB-IoT module	SIM7080G, supports CoAP-over-NB-IoT, low bandwidth, ultra-low power consumption
Embedded OS	ESP-IDF v5.0, official firmware development environment for ESP32
Security library	liboqs v0.8.0, supports CRYSTALS-Kyber, SPHINCS+, and multiple PQC algorithms
Application protocol	CoAP, well-suited for NB-IoT due to its low bandwidth requirements and lightweight protocol stack

This configuration, outlined explicitly in Table 2, emulates practical NB-IoT deployment environments. The evaluation particularly emphasizes optimizing digital signature generation and transmission efficiency, diverging from broader integration evaluations of previous research.

Performance metrics were meticulously collected through several measurement techniques:

- Execution latency at each operational stage is measured by $esp\_timer\_get\_time()$.
- Average current draw (in mA) tracked with a USB power meter.
- Dynamic memory usage is monitored using $heap\_caps\_get\_free\_size(MALLOC\_CAP\_8BIT)$ before and after cryptographic processes.
- Network packet size and transmission latency were analyzed using Wireshark.

These measurement methodologies ensured consistent and reproducible results across multiple experimental iterations.

5.2. Measured performance metrics

Table 3 enumerates the average execution time recorded across the cryptographic operation stages. The cumulative execution time is notably higher than traditional cryptographic schemes like ECDSA or RSA. Nevertheless, this increment is offset by the long-term security assurances provided by post-quantum algorithms. Furthermore, employing session reuse and optimized scheduling significantly diminishes the frequency of extensive signing operations, effectively moderating the performance impact on constrained NB-IoT networks.

Table 3: Average execution time per stage

Stage	Average execution time (ms)
Session Initialization	2.1 ms
Key Exchange (Kyber)	38.6 ms
Digital Signature (SPHINCS+)	92.3 ms
Data Encryption (ChaCha20-Poly1305)	2.1 ms
Total	133.0 ms

Table 4 illustrates the packet size overhead introduced by the incorporation of post-quantum cryptographic components.

Table 4: Packet Size Overhead Introduced by Post-Quantum Components

Component	Additional size (bytes)
Kyber Public Key	800 bytes
SPHINCS+ Signature	34,000 bytes
Poly1305 Authentication Tag	16 bytes
Total Overhead	∼34,816 bytes (∼34KB)

As observed from Table 4, deploying SPHINCS+ significantly inflates packet size, greatly exceeding the typical NB-IoT payload limits (∼1-2 KB), thus presenting a substantial challenge to network efficiency.

To mitigate these challenges, we employed techniques such as signature aggregation and Merkle tree pruning, achieving notable packet size reduction. Table 5 summarizes the efficacy of these optimization methods.

Table 5: Reduction achieved through optimization

Technique	Avg. signature size	% Reduction	Security level
Default SPHINCS+	34 KB	—	Category 4
Merkle Tree Pruning	18 KB	47%	Category 3
Signature Aggregation (5 nodes)	6 KB per node (total 30 KB for all)	60%	Same as default
Selective Signing (every 3 transmissions)	34 KB per session	66% effective reduction	Forward secrecy preserved

Among these, signature aggregation emerged as the most effective strategy in reducing packet overhead, closely followed by selective signing policy and Merkle tree pruning. Although none of these approaches fully eradicate the large signature sizes inherent to SPHINCS+, they offer practical solutions suitable for NB-IoT environments.

Additionally, employing ephemeral session keys enhances forward secrecy, thereby preserving the confidentiality of historical communications even in scenarios where future keys become compromised, a crucial security property for enduring IoT deployments.

6. DISCUSSION AND PRACTICAL CONSIDERATIONS

The deployment of post-quantum digital signatures like SPHINCS+ in Narrowband IoT (NB-IoT) systems introduces a unique set of practical challenges that must be carefully addressed to ensure real-world feasibility. Unlike classical cryptographic schemes such as ECDSA or RSA, which generate compact signatures (∼64 - 128bytes), SPHINCS+ produces metadata that can reach up to ∼34KB per transmission, significantly exceeding the typical payload limit (∼1 - 2KB) of NB-IoT networks [17]. This makes it essential to explore optimization strategies that reduce overhead while maintaining quantum resilience.

Our experimental results show that packet size expansion due to SPHINCS+ is substantial but manageable through targeted mitigation techniques:

Signature aggregation: Allows multiple devices to share a single signature block, reducing total network load by up to 60%.

Merkle tree pruning: Reduces signature size from ∼34 KB to ∼18 KB at the expense of lowering security category from NIST's Category 4 to Category 3, still acceptable for many industrial deployments.

Selective signing policy: Avoids unnecessary invocation of SPHINCS+, executing it only during critical transmissions or initial authentication phases. This strategy improves energy efficiency without compromising forward secrecy.

Offloading verification: Shifts the heavy computation of signature verification to gateways or cloud services, minimizing client-side overhead and enabling backward compatibility with legacy NB-IoT modules [11].

These findings align with prior work on lightweight PQC integration, particularly studies emphasizing deployability over theoretical novelty. However, unlike earlier efforts that primarily evaluated computational cost or memory footprint, our focus lies specifically on packet-level efficiency, a critical constraint in LPWAN environments where bandwidth is limited and data transmission intervals are long.

A notable trade-off between security level and packet size was identified in this study. While full-strength SPHINCS+-SHA256-128f offers the highest resistance to forgery attacks, its large signature size makes it unsuitable for frequent use in NB-IoT unless mitigation strategies are applied. On the other hand, compact variants like SPHINCS+-SHA256-128s offer reduced overhead and acceptable security for applications where firmware updates are rare and long-term confidentiality is not an immediate concern.

While the compact variant SPHINCS+-SHA256-128s reduces signature size by nearly 47%, it operates at NIST Security Category 3 (equivalent to AES-128) [22], as opposed to Category 4 (AES-192) [23] of the full variant. This reduction implies a lower resistance margin against future quantum attacks, particularly those leveraging Grover's algorithm or improved collision-finding techniques. For systems handling highly sensitive or long-lived data, such as national infrastructure logs, medical records, or legal evidence, the use of Category 3 may not meet regulatory or compliance requirements. However, for many industrial IoT deployments, such as smart metering, environmental monitoring, or utility sensor networks with data validity periods are under 10 - 15 years. Category 3 remains sufficient given current projections of quantum computing advancement. We recommend a risk-based deployment model: pruned variants should be used only when combined with additional mitigations such as short data retention policies, periodic re-authentication, or hybrid authentication schemes that layer hash-based signatures with lattice-based backups.

Another key observation is that offloading verification to gateway-level processing units significantly reduces the burden on constrained NB-IoT clients. This approach allows devices to send only public keys and message hashes, leaving full signature validation to backend infrastructure with greater computational capacity [11]. It also supports gradual migration toward post-quantum infrastructure. Organizations can begin verifying SPHINCS+ signatures at the gateway today, even if client-side firmware updates are delayed until later stages.

A significant practical limitation of signature aggregation is its reliance on synchronized communication among devices. In dynamic or event-driven NB-IoT networks, where transmissions are asynchronous or triggered by unpredictable events, achieving perfect synchronization is often infeasible. To enhance applicability, we propose a time-windowed aggregation mechanism: the gateway buffers incoming messages within a configurable time window (e.g., ±5 minutes around scheduled transmission times) and applies aggregation if multiple nodes transmit within that interval. If no co-transmissions occur, individual signatures are processed without penalty. This hybrid approach preserves the benefits of aggregation in scheduled deployments (e.g., utility meter reading) while maintaining flexibility for unsynchronized scenarios. Minimal coordination overhead is introduced since no handshake protocol is required, only loose temporal alignment based on network-wide time synchronization (e.g., via NITZ or GPS-derived timestamps).

Offloading verification introduces a trusted computing base at the edge, potentially expanding the attack surface if the gateway is compromised. A malicious actor gaining control of the gateway could accept forged packets without detection, undermining end-to-end authenticity. To mitigate this risk, we advocate for a defence-in-depth strategy: (i) securing the gateway with hardware-rooted trust (e.g., TPM, Secure Enclave, or TrustZone); (ii) enforcing mutual authentication between devices and the gateway using pre-shared keys or lightweight certificates; and (iii) logging all verification outcomes for anomaly detection and audit trails. Additionally, gateways should be designed with a higher update frequency than end devices, enabling timely patching and support for evolving PQC standards. This architectural separation allows constrained NB-IoT nodes to benefit from post-quantum security without bearing the full computational burden, provided that edge-layer trust is rigorously maintained.

Compared to earlier research such as Abbasi et al. [21] and Kumari et al. [24], our work addresses a specific scalability issue previously overlooked, the impact of PQC metadata on NB-IoT packet size. We build upon their findings by introducing concrete technical solutions that reduce overhead without compromising security guarantees.

Despite promising outcomes, several limitations remain:

- Aggregation requires synchronized communication, making it unsuitable for asynchronous or independent device transmissions.

- Tree pruning lowers the security level from NIST's Category 4 to Category 3, which may not meet regulatory requirements in highly sensitive environments.

- Selective signing policies must be carefully managed to avoid reducing message freshness below acceptable levels.
- Offloading increases dependency on backend infrastructure, which could introduce new vulnerabilities if gateways are compromised.

These constraints highlight the importance of tailoring post-quantum optimizations to specific deployment scenarios. For example:

- In smart metering applications with scheduled transmissions, signature aggregation is ideal.
- In remote sensing environments with low update frequency, tree pruning can be safely applied.
- In systems with high energy conservation needs, a selective signing policy proves most effective.
- In hybrid infrastructures with strong backend support, offloading verification provides the best balance between efficiency and security.

It is important to clarify that while our framework incorporates Kyber for key exchange and ChaCha20-Poly1305 for encryption to form a complete post-quantum secure channel, the optimization focus remains exclusively on reducing the packet overhead introduced by SPHINCS+ signatures, the primary source of size inflation in stateless hash-based PQC schemes. The contributions of this work lie in mitigating the most significant bottleneck in NB-IoT post-quantum deployment: large signature metadata. Other elements of the cryptographic stack, though essential for system completeness, contribute negligibly to packet expansion and are therefore treated as supporting components rather than optimization subjects.

Overall, our framework demonstrates that SPHINCS+ remains viable for NB-IoT deployments when combined with intelligent scheduling and packet-level optimization techniques. As standardization efforts by NIST and ETSI continue to evolve, transitioning toward post-quantum infrastructure is no longer a distant requirement, it is something we must act upon today, especially for systems with multi-year deployment cycles.

7. CONCLUSION AND FUTURE WORK

Unlike lattice-based schemes such as Dilithium or Falcon, which require complex state management and are vulnerable to timing attacks [14]. New proposals such as HQC and BIKE are also vulnerable to this risk [25, 26]; Our proposal uses SPHINCS+, the only standardized stateless hash-based digital signature scheme selected by NIST after the final round of its PQC standardization process [8]. While our system integrates Kyber for key exchange and ChaCha20-Poly1305 for data encryption to form a complete hybrid post-quantum secure channel, it is crucial to emphasize that the primary focus of this work lies exclusively on mitigating the dominant source of packet expansion: SPHINCS+ signatures. These additional components, though necessary for end-to-end security, contribute negligibly to payload inflation and are treated as supporting layers rather than optimization targets.

Compared to earlier studies such as Abbasi et al. [21] and Kumari et al. [24], our work introduces notable differences:

- A narrow, problem-driven scope centered on packet-level inefficiency, not just computational feasibility, aligns with real-world deployment constraints in LPWAN environments.
- Practical mitigation strategies specifically tailored for NB-IoT systems, where firmware updates are rare and energy conservation is critical.
- An empirical evaluation demonstrating tangible reductions in network load through system-level engineering, rather than algorithmic novelty alone.

Despite promising outcomes, several directions for further enhancement remain.

Time-windowed signature aggregation: To address the current limitation of requiring strict synchronization among devices, future work will explore time-windowed aggregation mechanisms. In this model, the gateway buffers incoming messages within a configurable time window (e.g., ±5 minutes) and applies aggregation if multiple nodes transmit within that interval. This approach enhances flexibility for asynchronous or event-driven deployments while preserving bandwidth savings in scheduled scenarios.

Trust-hardened gateways for offloaded verification: Since offloading shifts the trust boundary to the gateway, future research will investigate integrating hardware-rooted trust (e.g., TrustZone-M, Secure Enclave) into edge nodes to minimize risks of compromise. Additionally, we will develop lightweight logging and anomaly detection modules to monitor verification integrity at the gateway level.

Risk-adaptive pruning policies: We plan to develop context-aware deployment models that dynamically select between full-strength (Category 4) [23] and pruned (Category 3) [22] variants of SPHINCS+ based on application requirements. For instance, Category 4 will be enforced for long-term sensitive data (e.g., medical logs), while Category 3 may suffice for short-lived sensor readings. This risk-based strategy ensures an optimal balance between efficiency and regulatory compliance.

Extension to other LPWAN technologies: The proposed framework will be adapted to LoRaWAN and LTE-M, broadening its applicability beyond NB-IoT. Given their similar bandwidth constraints, these networks face analogous challenges with PQC metadata, making them natural candidates for our optimization techniques.

Deployment on smaller microcontrollers: We aim to test the feasibility of our methods on platforms with tighter resource limits, such as STM32WB and RP2040, to assess scalability under stricter memory and power budgets.

Integration with decentralized authentication models: Exploring lightweight blockchain-based attestation or zero-knowledge proofs could complement our current centralized verification model, offering enhanced resilience in distributed IoT architectures.

These development paths provide a roadmap for advancing hybrid frameworks that balance computational efficiency, energy conservation, and long-term security guarantees. As standardization efforts by NIST and ETSI accelerate, transitioning toward post-quantum infrastructure is no longer a distant goal. It is something we must act upon today.

Our approach does not seek to replace classical cryptographic schemes overnight, but rather to provide a realistic path toward gradual migration, one that balances backward compatibility with forward secrecy. We advocate for a phased transition model where PQC is introduced selectively during high-risk phases, while lightweight schemes continue to serve lower-priority functions until full replacement becomes viable.

This study contributes a deployable solution for system architects and firmware developers aiming to integrate post-quantum security into NB-IoT systems. By focusing on real-world implementation, empirical evaluation, and system-level integration, rather than theoretical novelty alone, this work lays the foundation for future research into deploying PQC-based frameworks in industrial IoT systems where backward compatibility and forward secrecy are essential.

ACKNOWLEDGMENTS

The authors wish to thank Thai Nguyen University of Information and Communication Technology. This work was supported in part by a grant number: DH2025-TN07-04.

REFERENCES

[1] N. Minh, D. Moldovyan, N. Moldovyan, A. Kostina, L. Minh, L. Huong, and N. Giang, “Post-quantum blind signature protocol on non-commutative algebras,” *Journal of Computer Science and Cybernetics*, vol. 37, no. 4, pp. 495–509, 2021. [Online]. Available: https://doi.org/10.15625/1813-9663/37/4/16023

[2] K. L. Dinh, L. G. Nguyen, T. B. Do, A. M. Alexandr, N. M. Dmitriy, and A. K. Anna, “Defining high-dimensional non-commutative algebras as carriers for post-quantum digital signature algorithms,” in *2024 1st International Conference On Cryptography And Information Security (VCRIS)*. IEEE, 2024, pp. 1–5. [Online]. Available: https://doi.org/10.1109/VCRIS63677.2024.10813386

[3] V. Lyubashevsky, L. Ducas, E. Kiltz, T. Lepoint, P. Schwabe, G. Seiler, D. Stehlé, and S. Bai, “Crystals-dilithium: Algorithm specifications and supporting documentation,” *NIST Post-Quantum Cryptography Standardization Round*, vol. 3, 2020. [Online]. Available: https://pq-crystals.org/dilithium/

[4] P.-A. Fouque, J. Hoffstein, P. Kirchner, V. Lyubashevsky, T. Pornin, T. Prest, T. Ricosset, G. Seiler, W. Whyte, and Z. Zhang, “Falcon: Fast-fourier lattice-based compact signatures over NTRU,” *Submission to the NIST's post-quantum cryptography standardization process*, vol. 36, no. 5, pp. 1–75, 2018. [Online]. Available: https://falcon-sign.info/

[5] F. Antognazza, A. Barenghi, G. Pelosi, and R. Susella, “A high efficiency hardware design for the post-quantum KEM HQC,” in *2024 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, 2024, pp. 431–441. [Online]. Available: https://doi.org/10.1109/HOST55342.2024.10545409

[6] R. Asif, “Post-quantum cryptosystems for internet-of-things: A survey on lattice-based algorithms,” *IoT*, vol. 2, no. 1, pp. 71–91, 2021. [Online]. Available: https://doi.org/10.3390/iot2010005

[7] R. S. Sinha, Y. Wei, and S.-H. Hwang, "A survey on LPWA technology: LoRa and NB-IoT," *Ict Express*, vol. 3, no. 1, pp. 14–21, 2017. [Online]. Available: https://doi.org/10.1016/j.icte.2017.03.004

[8] E. Gerjuoy, "Shor's factoring algorithm and modern cryptography an illustration of the capabilities inherent in quantum computers," *American Journal of Physics*, vol. 73, no. 6, pp. 521–540, 2005. [Online]. Available: https://doi.org/10.1119/1.1891170

[9] P. Ravi, A. Chattopadhyay, J. P. D'Anvers, and A. Baksi, "Side-channel and fault-injection attacks over lattice-based post-quantum schemes (kyber, dilithium): Survey and new results," *ACM Transactions on Embedded Computing Systems*, vol. 23, no. 2, pp. 1–54, 2024. [Online]. Available: https://doi.org/10.1145/3603170

[10] D. J. Bernstein, A. Hülsing, S. Kölbl, R. Niederhagen, J. Rijneveld, and P. Schwabe, "The SPHINCS+ signature framework," in *Proceedings of The 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019, pp. 2129–2146. [Online]. Available: https://doi.org/10.1145/3319535.3363229

[11] K. Bürstinghaus-Steinbach, C. Krauß, R. Niederhagen, and M. Schneider, "Post-quantum TLS on embedded systems: Integrating and evaluating kyber and SPHINCS+ with mbed TLS," in *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, 2020, pp. 841–852. [Online]. Available: https://doi.org/10.1145/3320269.3384725

[12] M. Chen, Y. Miao, Y. Hao, and K. Hwang, "Narrow band internet of things," *IEEE access*, vol. 5, pp. 20 557–20 577, 2017. [Online]. Available: https://doi.org/10.1109/ACCESS.2017.2751586

[13] D. J. Bernstein, D. Hopwood, A. Hülsing, T. Lange, R. Niederhagen, L. Papachristodoulou, M. Schneider, P. Schwabe, and Z. Wilcox-O'Hearn, "SPHINCS+: practical stateless hash-based signatures," in *Annual International Conference on The Theory and Applications of Cryptographic Techniques*. Springer, 2015, pp. 368–397. [Online]. Available: https://doi.org/10.1007/978-3-662-46800-5_15

[14] J. Jancar, M. Fourné, D. D. A. Braga, M. Sabt, P. Schwabe, G. Barthe, P.-A. Fouque, and Y. Acar, ""they're not that hard to mitigate": What cryptographic library developers think about timing attacks," in *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2022, pp. 632–649. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/SP46214.2022.9833713

[15] J. Lopez Valdivieso and R. Cumplido, "Design and implementation of hardware-software architecture based on hashes for SPHINCS+," *ACM Trans. Reconfigurable Technol. Syst.*, vol. 17, no. 4, Oct. 2024. [Online]. Available: https://doi.org/10.1145/3653459

[16] G. M. Zaverucha and D. R. Stinson, "Short one-time signatures," *Cryptology ePrint Archive*, 2010. [Online]. Available: https://eprint.iacr.org/2010/446

[17] M. Rana, Q. Mamun, and R. Islam, "Lightweight cryptography in IoT networks: A survey," *Future Generation Computer Systems*, vol. 129, pp. 77–89, 2022. [Online]. Available: https://doi.org/10.1016/j.future.2021.11.011

[18] N. Gupta, A. Jati, A. K. Chauhan, and A. Chattopadhyay, “PQC acceleration using gpus: Frodokem, newhope, and kyber,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 3, pp. 575–586, 2020. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/TPDS.2020.3025691

[19] D. O. Greconici, M. J. Kannwischer, and A. Sprenkels, “Compact dilithium implementations on Cortex-M3 and Cortex-M4,” *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp. 1–24, 2021. [Online]. Available: https://doi.org/10.46586/tches.v2021.i1.1-24

[20] Y. Zhou, K. Rajasekaran, and Q. Wang, “Exploring parallel implementation of sphincs+ using advanced vector extensions (avx) sets,” in *2025 26th International Symposium on Quality Electronic Design (ISQED).* IEEE, 2025, pp. 1–8. [Online]. Available: https://doi.org/10.1109/ISQED65160.2025.11014474

[21] M. Abbasi, F. Cardoso, P. Váz, J. Silva, and P. Martins, “A practical performance benchmark of post-quantum cryptography across heterogeneous computing environments,” *Cryptography*, vol. 9, no. 2, 2025. [Online]. Available: 10.3390/cryptography9020032

[22] G. M. D. Nist, “Module-lattice-based key-encapsulation mechanism standard,” National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep., 2024. [Online]. Available: https://doi.org/10.6028/NIST.FIPS.203

[23] ——, “Module-Lattice-Based digital signature standard,” National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep., 2024. [Online]. Available: https://doi.org/10.6028/NIST.FIPS.204

[24] S. Kumari, M. Singh, R. Singh, and H. Tewari, “Post-quantum cryptography techniques for secure communication in resource-constrained internet of things devices: A comprehensive survey,” *Software: Practice and Experience*, vol. 52, no. 10, pp. 2047–2076, 2022. [Online]. Available: https://doi.org/10.1002/spe.3121

[25] Q. Guo, C. Hlauschek, T. Johansson, N. Lahr, A. Nilsson, and R. L. Schröder, “Don’t reject this: Key-recovery timing attacks due to rejection-sampling in HQC and BIKE,” *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pp. 223–263, 2022. [Online]. Available: https://doi.org/10.46586/tches.v2022.i3.223-263

[26] J. P. D’Anvers, M. Tiepelt, F. Vercauteren, and I. Verbauwhede, “Timing attacks on error correcting codes in post-quantum schemes,” in *Proceedings of ACM Workshop on Theory of Implementation Security Workshop*, 2019, pp. 2–9. [Online]. Available: https://doi.org/10.1145/3338467.3358948

Received on July 13, 2025
Accepted on October 21, 2025