周报 向嘉豪(2025-11-24)

摘要: 本周完成了 ML-DSA 签名算法的深入分析与 NTT 优化策略研究工作。新增签名算法结构和计算特征章节，包含签名流程概述（wrapper 函数与内部签名程序）、Fiat-Shamir with Aborts 范式的拒绝采样机制、数论变换(NTT)正向与逆向算法伪代码。分析 ARM Cortex-M4 平台四个性能瓶颈（拒绝采样开销、NTT 计算主导、模约减开销、哈希函数调用），阐述四类优化策略（NTT 汇编优化、延迟模约减、Barrett/Montgomery 约减、预计算策略）。

下周计划: 1) 实现 NTT 汇编优化或延迟模约减策略 2) 论文 Results and Analysis 章节撰写

1 签名算法结构与计算特征分析

完成了 ML-DSA 签名算法结构和计算特征章节撰写，从算法流程、核心变换和计算复杂度三个维度分析签名操作的计算瓶颈。签名操作包含 wrapper 函数和内部签名程序两个层级，wrapper 函数处理上下文字符串验证（限制 255 字节）和随机数生成（32 字节），确定性变体使用全零随机数实现可重现签名。内部签名程序实现 Fiat-Shamir with Aborts 范式，展开私钥后进入拒绝采样循环：每次迭代采样掩蔽向量$\boldsymbol{y}$，计算承诺$\boldsymbol{w} = \boldsymbol{A} \cdot \boldsymbol{y}$，从承诺哈希推导挑战多项式$c$，评估响应向量$\boldsymbol{z} = \boldsymbol{y} + c \cdot \boldsymbol{s}_1$。响应向量需通过边界检验（阈值$\gamma_1 - \beta$），超限则重启迭代。

数论变换(NTT)构成 ML-DSA 签名的计算核心，实现环$R_q = \frac{\mathbb{Z}_q[X]}{X^{256}+1}$上的高效多项式乘法，其中$q = 8380417$。正向 NTT 采用 Cooley-Tukey 蝶形结构，将多项式从系数表示变换到 512 次本原单位根$\zeta = 1753$幂次处的求值表示，通过 8 个阶段（$\log_2 256 = 8$）完成变换，每阶段减半步长。逆向 NTT 采用 Gentleman-Sande 蝶形结构配合取反旋转因子，最终乘以缩放因子$f = 8347681 \equiv 256^{-1} \bmod q$归一化输出。单次 NTT 执行需要$256 \times 8 = 2048$次蝶形操作，每次蝶形涉及一次模乘和两次模加减，参考实现在每次算术操作后执行模约减。签名流程每次迭代调用多次 NTT/INTT：掩蔽向量$\boldsymbol{y}$的正向 NTT、NTT 域内矩阵向量乘法$\boldsymbol{A} \cdot \hat{\boldsymbol{y}}$、承诺计算的逆向 NTT，结合拒绝采样迭代（期望次数 4.25/5.1/3.85），NTT 操作在 ARM Cortex-M4 平台占签名总计算成本的 60-70%。

2 性能瓶颈与优化策略

完成了 ML-DSA 签名在资源受限平台的性能瓶颈分析与优化策略阐述。性能剖析识别四个主要瓶颈：拒绝采样开销源于 Fiat-Shamir with Aborts 范式的迭代签名尝试，ML-DSA-44/65/87 的期望迭代次数分别为 4.25、5.1 和 3.85，最坏情况可超过 20 次迭代；NTT 计算主导地位明确，参考 NTT 实现在 168 MHz ARM Cortex-M4 上每次变换消耗 250-300 万周期；模约减开销占 NTT 计算成本的约 40%，参考实现采用基于除法的模约减在每次蝶形操作后执行；SHAKE-256 哈希函数调用用于矩阵展开、挑战推导和消息哈希，消耗 15-20% 签名计算量。

针对识别的性能瓶颈，阐述四类优化策略作为后量子嵌入式密码学的研究方向。NTT 汇编优化通过手工优化的 ARM 汇编实现指令级并行、寄存器分配优化和流水线调度，利用 UMULL 指令实现$32 \times 32 \to 64$位乘法和条件执行消除分支，相对编译 C 实现达到 20-30% 延迟降低。延迟模约减策略通过跨多个蝶形操作推迟模约减降低约减频率，维持中间值在扩展边界内（系数$< 2q$而非$< q$），经过溢出分析后实现 15-25% NTT 延迟改进。Barrett 和 Montgomery 约减以基于乘法的技术替代基于除法的模约减，Barrett 约减预计算$\mu = \left\lfloor \frac{2^{48}}{q} \right\rfloor$通过乘法和移位操作实现约减，Montgomery 约减提供高效的融合乘-约减操作，在 ARM Cortex-M4 上实现 25-35% 约减开销改进。预计算策略将秘密向量$\boldsymbol{s}_1$、$\boldsymbol{s}_2$存储为 NTT 表示消除逐次签名的 NTT 变换，以 10-13 KB

额外 Flash 存储换取 20-25% 签名延迟降低，旋转因子预计算（512 项，2 KB）消除运行时单位根幂次计算。上述优化技术非互斥，组合实现相对参考实现可达 40-50% 累积改进，提升 ML-DSA 在资源受限 IoT 部署的可行性。