

基于 Polar 码改进的 RLCE 公钥加密方案

李喆¹, 韩益亮^{1,2}, 李鱼¹

(1. 武警工程大学密码工程学院, 陕西 西安 710086;

2. 武警部队密码与信息安全保密重点实验室, 陕西 西安 710086)

摘要: 针对 PolarRLCE 方案不具有语义安全, 易受到自适应选择密文攻击 (IND-CCA2, adaptively chosen ciphertext attacks) 的缺点。在 RLCE (random linear code encryption) 方案的基础上, 利用 RLCE 方案的结构和 Polar 码的极化性质, 将 Polar 码作为方案的底层编码, 通过 RLCEspad 消息填充的方法, 采用普通编码对密文进行编码, 提出一种具有语义安全、可以达到 IND-CCA2 安全的改进的 RLCE 公钥加密方案。改进后的方案将公钥矩阵转为系统矩阵, 减小了公钥存储空间; 对部分私钥进行预计算, 减小了私钥存储空间。通过分析, 所提方案未改变 PolarRLCE 方案的结构, 可以抵抗针对汉明循环码的结构化等攻击。在 128 bit 安全级别, 相较于 HermitianRLCE 方案、GRSRLCE 方案和 GoppaMcEliece 方案, 所提方案的公钥尺寸分别减少了 4%、46.5%、47.9%。

关键词: Polar 码; 随机线性码加密; 消息填充; 自适应选择密文攻击; 结构化攻击

中图分类号: TP309.7

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2020068

Improved RLCE public key encryption scheme based on Polar codes

LI Zhe¹, HAN Yiliang^{1,2}, LI Yu¹

1. College of Cryptographic Engineering, Engineering University of PAP, Xi'an 710086, China

2. Key Laboratory of PAP for Cryptology and Information Security, Xi'an 710086, China

Abstract: For PolarRLCE has no semantic security and be vulnerable to adaptively chosen ciphertext attacks. Based on random linear code encryption scheme, the structure of the RLCE scheme and Polar code polarization properties were used to make the Polar code as underlying encoding scheme, and through the method of RLCEspad message padding, a kind of improved public key encryption scheme which considers semantic security and can resistant to adaptively chosen ciphertext attacks was proposed. The improved scheme transformed the public key matrix into a

收稿日期: 2020-01-03; 修回日期: 2020-03-21

通信作者: 韩益亮, hanyil@163.com

基金项目: 国家自然科学基金 (61572521); 武警工程大学科研创新团队科学基金 (KYTD201805)

Foundation Items: The National Natural Science Foundation of China (61572521), The Scientific Foundation of the Scientific Research and Innovation Team of Engineering University of PAP (KYTD201805)

论文引用格式: 李喆, 韩益亮, 李鱼. 基于 Polar 码改进的 RLCE 公钥加密方案[J]. 网络与信息安全学报, 2020, 6(5): 110-118.

LI Z, HAN Y L, LI Y. Improved RLCE public key encryption scheme based on polar codes[J]. Chinese Journal of Network and Information Security, 2020, 6(5): 110-118.

system matrix and reduces the storage space of the public key. Some pre-computation for private key are estimated to reduce the storage space of private key. Through analysis, this scheme does not change the structure of PolarRLCE scheme, which can resist structural attacks against hamming quasi-cyclic codes. At the 128 bit security level, compared with HermitianRLCE scheme, GRSRLCE scheme and GoppaMcEliece scheme, the public key size of this scheme decreased by 4%, 46.5% and 47.9% respectively.

Key words: Polar code, random linear code encryption, message padding, adaptively chosen ciphertext attacks, structural attacks

1 引言

量子计算机技术的飞速发展, 将深刻影响当前和未来的社会。银行交易、无人驾驶、保密通信等与人们息息相关的各方面的安全性依赖于密码学。Shor^[1]提出的量子算法, 极大影响了基于大整数分解困难问题的 RSA 等^[2]公钥密码方案; Grover^[3]提出的量子算法, 极大影响了 AES 等^[4]对称密码方案。Shor 算法和 Grover 算法的提出给经典密码方案的安全性带来严峻的挑战。

为了应对量子计算机对基于经典数学困难问题的密码方案的威胁, 各国纷纷寻求能够抵抗量子计算机攻击的新型密码方案, 即后量子密码或抗量子密码 (PQC, post-quantum cryptography)。美国 NIST (National Institute of Standards and Technology) 在 2012 年启动后量子算法的标准化项目, 2016 年面向全球征集后量子密码算法, 2019 年征集算法完成第二轮遴选^[5]。欧洲电信标准化协会 (ETSI, European Telecommunications Standards Institute) 开展量子安全会议以此推动后量子密码算法的征集工作, 逐步推进全球后量子密码算法项目 PQCRYPTO 和后量子密码算法应用项目 SAFCRYPT^[6]。在以上各国或地区制定后量子密码算法标准的同时, 我国也同步推进后量子密码算法设计竞赛活动, 制定后量子算法标准, 截至 2019 年, 后量子密码标准化已进入第二轮评选。后量子密码方案主要包括 5 种^[7], 5 种密码方案各有其独特的应用范围。目前, 基于编码的方案和基于格的方案成为研究的热点。基于编码的密码方案相较于其他类型的后量子密码方案, 更适合构造加密方案, 在后量子密码加密方案中, 基于编码的加密方案具有良好的研究前景。

基于编码的密码方案最早可追溯到 1978 年

的 McEliece 方案^[8], 该方案最初使用 Goppa 码作为底层编码, 该方案在适当的参数选择下仍然是安全的, 并进入 PQC 第二轮征集算法。基于 Goppa 码的 McEliece 方案的安全性是基于一般线性码的译码困难问题, 可以规约到 NPC (non-deterministic polynomial complete) 问题, 具有抵抗量子计算机攻击的特点, 众多密码学者开始研究此方案。1986 年, Niederreiter^[9]利用 McEliece 方案的对偶形式构造了一种加密方案, 即 Niederreiter 方案。McEliece 方案和 Niederreiter 方案均存在密钥尺寸过大的缺点, 在实际场景中难以应用。针对 McEliece 方案实用化差的特点, 密码学者通过利用其他结构更加紧凑的编码作为底层编码以减小密钥尺寸, 然而构造的变形方案存在安全性问题, 易受到结构化攻击^[10]。如何在保证密码方案安全性的基础上, 同时构造较小密钥尺寸的密码方案成为基于编码的密码方案亟待解决的问题。

2016 年, Wang^[11]通过在生成矩阵的每一列中插入随机列构造了线性随机码公钥加密 (RLCE, random linear code encryption) 方案, 该方案进入 PQC 第一轮征集算法评选。不同于其他利用结构紧凑的编码的密码方案, 该方案不依赖于任何底层编码的结构。RLCE 方案通过插入随机列使方案达到随机性, 同时方案的安全性依赖于线性随机码译码的 NPC 问题, 避免了因底层编码结构引入的结构化攻击。2017 年, Wang^[12]针对 RLCE 方案存在的潜在攻击, 进一步研究 RLCE 方案的填充方法, 改进后的 RLCE 方案减小了密钥尺寸, 提高了加解密性能, 达到了 IND-CCA2 安全性。Matthews^[13]在 2019 年 CBC (Code-Based Cryptography) 会议上提出用 Hermitian (埃尔米特) 码作为 RLCE 方案的底层编码, 在 128 bit 安全级别和 256 bit 安全级别的情况下, 相较于

GRSRLCE 加密方案、HermitianRLCE 加密方案的公钥尺寸分别减少了 80%、72%。Liu 等^[14]在 2019 年 A2C (Algebra, Codes and Cryptology) 会议上提出利用 Polar 码作为 RLCE 方案的底层编码, 通过利用 Polar 码的极化性质, 降低了加解密复杂度, 在正确参数选择下, 可以达到预期的安全级别, 同时密钥尺寸保持在合理的范围内。然而, Liu 在文献[14]提出的方案没有考虑语义安全, 易受到自适应选择密文攻击。

本文在 RLCE 方案和 PolarRLCE 方案的基础上, 利用 Polar 码的极化性质, 通过 RLCEspad 消息填充, 提出一种具有 IND-CCA2 安全的基于 Polar 码改进的 RLCE 公钥加密方案。改进的方案利用 Arikan^[15]提出的 Polar 码作为底层编码, 在信息比特列中插入随机列, 选择随机非奇异矩阵对随机列进行混合, 选择等价类较多的置换矩阵对生成矩阵进行扰乱, 利用消息填充的方法, 使用普通编码对密文进行编码。改进后的方案没有改变 RLCE 方案的结构, 仍然具有随机性, 可以抵抗 McEliece 方案存在的结构化攻击。此外, 本文采用系统矩阵的形式, 对部分私钥进行预计算, 减少了密钥存储空间。

2 基础知识

2.1 Polar 码编码理论

根据 Shannon 阐述的信道容量概念, Polar 码是至今唯一可以在理论上证明达到这一容量的编码。Polar 码具有良好的性能, 加解密复杂度低, 计算复杂度为 $O(n \log n)$ 。Polar 码由于其独特的极化性质、较低的加解密复杂度, 成为众多通信领域学者和密码领域学者研究的焦点。目前, 利用 Polar 码在物理层的极化性质, 许多学者在信息安全保密领域取得突破^[16]。华为利用 Polar 码信道容量大、传输速率高、安全性强的优点, 将 Polar 码作为 5G 时代无线通信的编码^[17]。

Polar 码相较于 Goppa 码具有更好的纠错能力, 较低的加解密复杂度, 较多的等价类码族。拥有以上优点的 Polar 码应用到编码密码方案中, 将具有合理的密钥尺寸, 更加适合构造基于 Polar 码的加密方案, 基于 Polar 码的加密方案具有良好的应用前景。Shrestha 等^[18]将 Polar 码应用到密

码方案中。随着 Polar 码理论的进一步发展, Hooshmand 等^[19]进一步研究 Polar 码在密码方案中的应用, 完善基于 Polar 码的密码方案。

定义 1 对称信道

Polar 码可以实现任何二进制输入的离散无记忆性信道的 (B-DMC, binary input discrete memoryless channel) 容量, 如二进制擦除信道 (BEC) 和二进制对称信道 (BSC)。下面阐述二进制对称信道相关定义。定义 X 是输入符号集, Y 是输出符号集, 转移概率为 $W(y|x)$, $x \in X$, $y \in Y$ 。信道 W 经过 N 次极化后, 信道可以表示为 W^N , 则信道 $W^N: X^N \rightarrow Y^N$ 的转移概率为

$$W^N(y_1^N | x_1^N) = \prod_{i=1}^N W(y_i | x_i) \quad (1)$$

关于 B-DMC 的性能, 可以通过两个参数进行评价: 对称容量 $I(W)$ 和巴氏参数 $Z(W)$ 。下面对两个参数相关定义进行阐述。

(1) 对称容量 (symmetric capacity)

$$I(W) \triangleq \sum_{y \in Y} \sum_{x \in X} \frac{1}{2} W(y|x) \log \frac{W(y|x)}{\frac{1}{2} W(y|0) + \frac{1}{2} W(y|1)} \quad (2)$$

(2) 巴氏参数 (Bhattacharyya parameter)

$$Z(W) \triangleq \sum_{y \in Y} \sqrt{W(y|0)W(y|1)} \quad (3)$$

$I(W) \in [0, 1]$ 用于评估信道速率, $Z(W) \in [0, 1]$ 用于评估信道可靠性。 $I(W)$ 与 $Z(W)$ 具有如下的关系: 当且仅当 $Z(W) \approx 0$ 时, $I(W) \approx 1$; 当且仅当 $Z(W) \approx 1$ 时, $I(W) \approx 0$ 。

定义 2 信道极化

信道极化分为两个阶段: 信道融合阶段 (channel combining) 和信道分化 (channel splitting) 阶段。

(1) 信道融合阶段描述

步骤 1 $W^N: X^N \rightarrow Y^N$, N 为 2 的 n 次幂 $N=2^n$, $n \geq 0$ 。其中, X 表示输入符号集合, Y 表示输出符号集合, W^N 表示信道 W 的 N 次使用后的信道。

步骤 2 $W_N(y_1^N | u_1^N) = W^N(y_1^N | u_1^N G_N)$, $y_i^N \in Y^N$, $u_i^N \in X^N$, G_N 为 N 维生成矩阵, $N=2^n$ 。其中, u_i^N 为原始比特输入序列; x_i^N 为通过生成矩阵编码后的比特序列。

(2) 信道分化阶段描述

步骤 1 将信道融合构成的复合信道 W_N 分裂为 N 个二进制输入的坐标信道。如下所示。

$$W_N: X \rightarrow Y^N \times X^{i-1}, 1 \leq i \leq N$$

$$\text{步骤 2 } W_N(y_1^N, u_1^{i-1} | u_i) \triangleq \sum_{u_i^N \in X^{N-i}} \frac{1}{2^{N-i}} W_N(y_1^N | u_1^N).$$

其中 (y_1^N, u_1^N) 表示 W_N^i 的输出, u_i 表示 W_N^i 的输入。

定义 3 极化码编码

极化码编码基本原理: 在信道巴氏参数 $Z(W_N)$ 趋近于 0, 或在信道对称容量 $I(W_N)$ 趋近于 1 的信道 W_N^0 上传输信息比特序列。

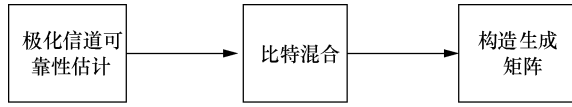


图 1 极化编码步骤
Figure 1 Polar coding steps

(1) 极化信道可靠性估计

对于 BEC, Arikan 给出的方法是计算巴氏参数 $Z(W_N^i)$

$$Z(W_N^i) = \sum_{y_1^N \in Y^N} \sum_{u_1^{i-1} \in X^{i-1}} \sqrt{W_N^i(y_1^N, u_1^{i-1} | 0) W_N^i(y_1^N, u_1^{i-1} | 1)} \quad (4)$$

巴氏参数 $Z(W_N^i)$ 越小, 即对称容量 $I(W_N^i)$ 越大, 说明选择的信道越适合传输信息比特; 反之, 巴氏参数 $Z(W_N^i)$ 越大, 即对称容量 $I(W_N^i)$ 越小, 说明信道越适合传输固定比特 (冻结比特)。

(2) 比特混合

可靠性较高的 K 个分裂子信道传输信息比特序列, 而可靠性较低的分裂子信道传输冻结比特。通过比特混合后, 输出的比特为编码的最初比特 u_1^N 。对于 BEC 来说, 选择巴氏参数 $Z(W_N^0)$ 最小的 K 个子信道传输信息比特。

(3) 构造生成矩阵

首先定义 Kronecker 矩阵乘法为

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} \quad (5)$$

定义矩阵 F 为 $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, 则生成矩阵

$G_N = B_N F^{\otimes n}$, 其中 $F^{\otimes n}$ 表示对矩阵 F 的 n 次克罗内克积, $F^{\otimes n} = F \otimes F^{\otimes (n-1)}$, $F^{\otimes n} F^{\otimes n} = I_N$ 。

B_N 是比特反转排序矩阵, 用来对比特重排操作。 B_N 的递归式定义为

$$B_N = R_N (I_2 \otimes B_{N/2}) \quad (6)$$

其中, I_2 为 2 维单位矩阵, $B_2 = I_2$; R_N 为置换矩阵, 用来分离输入序列中的奇偶元素。

2.2 RLCE 加密方案

(1) RLCE 密钥生成过程如算法 1 所示。

算法 1 RLCE 密钥生成

输入 $(n, k, d, t, w), n, k, d, t > 0, w \in \{1, 2, \dots, n\}$,

$k+1 \geq d \geq 2t+1$ 。

输出 公钥 G^{pub} , 私钥 (S, G_1, P, A) 。

1) 选择长度为 n , 维度为 k 的 $k \times n$ 生成矩阵 G 。

2) 随机选择 w 个 $k \times 1$ 的列向量 $r_1, r_2, \dots, r_w$, 将 w 个列向量随机插入生成矩阵 G 中, 得到 $k \times (n+w)$ 的生成矩阵 G_1 , $G_1 = (g_1, \dots, g_{n-w}, g_{n-w+1}, r_1, \dots, g_n, r_w)$ 。

3) 为了混合选择的列, 选择 w 个 2×2 随机非奇异矩阵 $A_1, A_2, \dots, A_w$ 。

4) 定义 $A = [1, \dots, 1, A_1, A_2, \dots, A_w]$ 是 $(n+w) \times (n+w)$ 的非奇异矩阵。

5) S 为 $k \times k$ 的非奇异矩阵, P 为 $(n+w) \times (n+w)$ 的置换矩阵。

6) 输出 $k \times (n+w)$ 的公钥 $G^{\text{pub}} = SG_1AP$, 私钥 (S, G_1, A, P) 。

(2) RLCE 加密过程如算法 2 所示。

算法 2 RLCE 加密

输入 公钥 G^{pub} , 明文 $m \in F_2^k$, 随机错误向量 $e \in F_2^{n+w}$ 。

输出 密文 $c \in F_2^{n+w}$ 。

$$c = mG^{\text{pub}} \oplus e, w_H(e) \leq t.$$

(3) RLCE 解密过程如算法 3 所示。

算法 3 RLCE 解密

输入 密文 $c \in F_2^{n+w}$, 私钥 (S, G_1, A, P) 。

输出 明文 $m \in F_2^k$, 译码错误标识 $\perp$ 。

1) $cP^{-1}A^{-1} = mSG_1 \oplus eP^{-1}A^{-1} = (c'_1, c'_2, \dots, c'_{n+w})$ 。

2) 从长度为 $n+w$ 的向量 $cP^{-1}A^{-1}$ 中删除 w 个

列向量, 得到长度为 n 的 $c' = (c'_1, c'_2, \dots, c'_{n-w+1}, c'_{n-w+3}, c'_{n-w+5}, \dots, c'_{n+w-1})$ 。

3) $c' = mSG_1 \oplus e'$, $e' \in F_2^n$, $w_H(e') \leq t$ 。

4) 利用译码算法, 计算 $m' = mS$, $m = m'S^{-1}$ 。

5) 计算汉明重量 $w = wt(c - mG_1)$, 假如 $w \leq t$, 输出明文 m ; 否则, 输出 $\perp$ 。

2.3 信息编码

为了抵抗编码密码方案存在的已知攻击, 可以通过消息编码 (消息填充) 的方法保证密码方案的安全性。通常, 在密文中有 3 种方法可以对消息进行编码^[12]。

(1) 基本编码 (basic encoding)

在向量 $m \in GF_2^k$ 中编码信息, 密文为 $c = mG + e$ 。在这种情况下, 可以在每个密文中编码长度为 $mLen = mk$ 位信息。

(2) 普通编码 (medium encoding)

除了基本编码之外, 更多的消息被编码在 e 的非零项中。在这种情况下, 可以在每个密文中编码长度为 $mLen = m(k+t)$ 位信息。严格地说, 因为 $e_{i_j} \neq 0$, 编码信息小于 $m(k+t)$ 位。

(3) 高级编码 (advanced encoding)

除了普通编码, 更多的消息被编码在 e 的非零项中。非零 e 共有 $\binom{n+w}{t}$ 种可能, 在这种情况下, 可以在每个密文中编码长度为 $mLen = m(k+t) + \left\lfloor \log_2 \binom{n+w}{t} \right\rfloor$ bit 的信息。

接下来概述常见的 5 种消息填充类型。

1) Pointcheval padding^[20]

$$c = \text{Enc}(G, r_1, H_1(m \| r_2)) \| (H_2(r_1) \oplus (m \| r_2)) \quad (7)$$

2) Fujisak-Okamoto padding^[21]

$$c = \text{Enc}(G, r_1, H_1(m \| r_1)) \| (H_2(r_1) \oplus m) \quad (8)$$

3) Kobara-Imai's α -padding^[22]

$$c = \text{Enc}(G, y_1, H_1(m \| r_1)) \| y_2(H_2(r_1) \oplus m), \quad (9)$$

$$y_1 \| y_2 = H_2(H_1(m \| r_1) \oplus (m \| r_1))$$

4) Kobara-Imai's β -padding

$$c = y_1 \| \text{Enc}(G, y_2, H_1(r_1)), \quad (10)$$

$$y_1 \| y_2 = (r \oplus H_1(H_2(r) \oplus m)) \oplus (H_2(r) \oplus m)$$

5) Kobara-Imai's γ -padding

$$c = y_3 \| \text{Enc}(G, y_1, y_2) \quad (11)$$

$$y_3 \| y_2 \| y_1 = (r \oplus H_1(H_2(r) \oplus (m \| \text{const}))) \| (H_2(r) \oplus (m \| \text{const})) \quad (12)$$

其中, H_1 、 H_2 是随机预言机模型 (也可以是伪随机位生成器或者哈希函数) 输出固定长度的伪随机比特串; r_1 、 r_2 是随机选择的适当长度的比特串。

3 本文的加密方案

3.1 改进的 RLCE 公钥加密方案

本文提出的基于 Polar 码改进的 RLCE 公钥加密方案是在 Wang 提出的 RLCE 方案和 Liu 提出的基于 Polar 码的 RLCE 方案的基础上, 针对 Liu 提出的 PolarRLCE 方案密钥尺寸较大且不具有 IND-CCA2 安全的缺点, 提出的一种改进方案。本文利用 Polar 码的极化性质和译码复杂度低的优点, 将 Polar 码作为 RLCE 方案的底层编码, 改进后的方案采用 Wang 的 RLCE 方案的结构, 没有改变 RLCE 方案的具体形式。将 Wang 的 RLCE 方案的 IND-CCA2 形式化安全证明应用到 Liu 的基于 Polar 码的编码方案中, 对 PolarRLCE 方案进行消息填充, 对每个密文进行语义安全普通编码, 并将公钥矩阵转化为系统矩阵, 对部分私钥采用预计算^[23], 以减少密钥存储空间。本文方案和 RLCE 方案类似, 包括密钥生成 (polarRLCE.KeySetup)、加密 (polarRLCE.Enc)、解密 (polarRLCE.Dec)。

(1) 密钥生成

密钥生成过程如下。

1) 参数选择: (n, k, d, t, w) , $n, k, d, t > 0$, $w \in \{1, 2, \dots, n\}$, $k+1 \geq d \geq 2t+1$ 。

2) G : 长度为 n , 维度为 k 的 Polar 码的 $k \times n$ 生成矩阵, 译码算法可以纠出至少 t 个错误。

3) G_1 : 长度为 $n+w$, 维度为 k 的 Polar 码的 $k \times (n+w)$ 的生成矩阵, 可以将随机选择的 w 个 $k \times 1$ 的列向量 $r_1, r_2, \dots, r_w$ 随机插入生成矩阵 G 得到 $G_1 = (g_1, \dots, g_{n-w}, g_{n-w+1}, r_1, \dots, g_n, r_w)$ 。

4) A : $(n+w) \times (n+w)$ 的非奇异矩阵

$$A = \begin{pmatrix} I_{n-w} & \cdots & 0 \\ \vdots & A_1 & \vdots \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_w \end{pmatrix}$$

其中, $A_1 = \begin{pmatrix} a_{1,11} & a_{1,12} \\ a_{1,21} & a_{1,22} \end{pmatrix}, \dots, A_w = \begin{pmatrix} a_{w,11} & a_{w,12} \\ a_{w,21} & a_{w,22} \end{pmatrix} \in F_2^{2 \times 2}$,

$A_1, A_2, \dots, A_w$ 是 2×2 随机非奇异矩阵。

5) S : $k \times k$ 阶的非奇异矩阵。

6) P : $(n+w) \times (n+w)$ 阶的置换矩阵。

7) G^{pub} : $k \times (n+w)$ 的公钥 $G^{\text{pub}} = SG_1AP$ 。

公钥 $G^{\text{pub}} = SG_1AP$, 私钥 (S, G_1, A, P) 。

(2) 加密

发送方对明文进行加密过程如下。

1) 发送方选择明文 $m \in F_2^k$, 随机选择错误向量 $e = [e_1, \dots, e_{n+w}] \in F_2^{n+w}$, 错误向量 e 的汉明重量至多为 t , 即 $wt_H(e) \leq t$ 。

2) 发送方利用接收方的公钥 G^{pub} 对明文 m 进行加密, 得到密文 $c \in F_2^{n+w}$ 。

$$c = mG^{\text{pub}} \oplus e \quad (13)$$

(3) 解密

接收方对密文进行解密过程如下。

1) 接收方收到密文 c , 利用自己的私钥, 对密文 c 右乘 $P^{-1}A^{-1}$, 得到

$$cP^{-1}A^{-1} = mSG_1 \oplus eP^{-1}A^{-1} = (c'_1, c'_2, \dots, c'_{n+w}) \quad (14)$$

2) 接收方从长度为 $n+w$ 的向量 $(c'_1, c'_2, \dots, c'_{n+w})$ 中删除 w 个冻结比特列, 得到长度为 n 的向量 $c' = (c'_1, c'_2, \dots, c'_{n-w+1}, c'_{n-w+3}, c'_{n-w+5}, \dots, c'_{n+w-1})$, $c' = mSG_1 \oplus e'$ (其中, $e' = eP^{-1}A^{-1}$), 此时的 $w_H(e') \leq t$ 。

3) 接收方利用 Polar 码的 SC 译码算法, 对 c' 进行译码, 得到明文 m 。

4) 接收方计算汉明重量 $wt = wt(c - mG_1)$, 假如 $wt \leq t$, 输出明文 m ; 否则, 输出 $\perp$ 。

3.2 将公钥矩阵转换为系统矩阵

令 $mSG_1 = (d_1, \dots, d_n)$, $c_d = (d'_1, \dots, d'_{n+w}) = (d_1, d_2, \dots, \perp, \dots, d_n, \perp)P$ 是长度为 $n+w$ 的向量。

对于每个 $i < k$, 当 $j < n-w$, $d'_i = d_j$, 令 $m_i = d_j$, 当 $m = (m_1, \dots, m_k)$ 。

$I_R = \{i : m_i\}$ (m_i 译码成功); $\bar{I}_R = \{1, \dots, k\} \setminus I_R$

(m_i 译码失败)。

假设 $|\bar{I}_R| = u$, 当 $i \in \bar{I}_R$ 时, 很容易得到剩下的 u 个 m_i 信息。对于每个 i_j , $i_1, \dots, i_u \geq k$ 为索引集, 当 $i < n-w$, $d'_i = d_i$ 。可以通过求下列方程组, 恢复出 $\bar{I}_R$ 中剩余的带有索引的消息。其中, $g_{i_1}, \dots, g_{i_u}$ 对应于公钥的每一列。

$$m[g_{i_1}, \dots, g_{i_u}] = [d'_{i_1}, \dots, d'_{i_u}] \quad (15)$$

P 为置换矩阵, 恢复出在 mP 中前 $k-u$ 个 $m_i (i \in I_R)$ 信息。通过式(15)计算得到

$$mPP^{-1}[g_{i_1}, \dots, g_{i_u}] = (m_{I_R}, m_{\bar{I}_R})P^{-1}[g_{i_1}, \dots, g_{i_u}] = [d'_{i_1}, \dots, d'_{i_u}] \quad (16)$$

其中, m_{I_R} 是在 I_R 带有索引的消息列表。

令 $P^{-1}[g_{i_1}, \dots, g_{i_u}] = \begin{pmatrix} V \\ W \end{pmatrix}$, 其中, V 是 $(k-u) \times u$ 阶

矩阵, W 是 $u \times u$ 阶矩阵。

$$m_{I_R}W + m_{\bar{I}_R}V = [d'_{i_1}, \dots, d'_{i_u}] \quad (17)$$

可以预计算 W^{-1} 。

$$m_{\bar{I}_R} = ([d'_{i_1}, \dots, d'_{i_u}] - m_{I_R}V)W^{-1} \quad (18)$$

通过将公钥矩阵转换为系统矩阵, 可以把 V, W^{-1} 作为私钥存储, 不再存储私钥 S , 以降低密钥存储空间。

3.3 消息填充

Liu 提出的 PolarRLCE 方案不具有 IND-CCA2 安全, 本文采用 Wang 提出的适合加密较短信息的 RLCEspad 消息填充的方法, 使改进的方案达到 IND-CCA2 安全。对信息进行填充后, 信息转化为 RLCE 方案中的信息位或其他信息, 采用普通编码, 填充的信息可以被编码在 e 的非零项中。

首先明确相关变量定义 RLCEspad ($mLen, k_1, k_2, k_3, t, m, r$)

$$(1) k_1 + k_2 + k_3 = \left\lceil \frac{mLen}{8} \right\rceil, v = 8(k_1 + k_2 + k_3) - mLen。$$

(2) H_1, H_2, H_3 是随机谕言机模型, 可以分别将任意长度的二进制比特串转化为长度为 $k_2, k_1 + k_2, k_3$ 字节的输出比特串。

(3) $m \in \{0, 1\}^{8k_1}$ 是需要填充的信息, 当 $r_1 \in$

$\{0,1\}^{8k_3-v}$, $r = r_1 \| 1^v$ 是随机选择的二进制比特串。

RLCE 填充过程如算法 4 所示。

算法 4 RLCEspad

输入 $mLen, t, m, r, H_1, H_2, H_3$ 。

输出 $(y_1, e_1), c$ 。

1) 计算 $v = 8(k_1 + k_2 + k_3) - mLen$ 。

2) 计算 $e_1 = l_1 \| l_2 \cdots \| l_t$, $l_i \in \{0,1\}^{16t}$, 其中 $m \in \{0,1\}^{8k_1}$, $r_1 \in \{0,1\}^{8k_3-v}$, $r = r_1 \| 1^v$ 。

3) 普通编码: $y = ((m \| H_1(m, r, e_1)) \oplus H_2(r, e_1)) \| H_3(m \| H_1(m, r, e_1)) \oplus H_2(r, e_1))$ 。

4) 将 y 转化为 $(y_1, e_1) \in GF_2^{k+t}$, 输出 (y_1, e_1) , 得到密文 $c = y_1 G + e$ 。

4 安全性分析

(1) 穷举攻击

穷举攻击是通过一一穷举可能存在的正确私钥, 从而恢复明文, 达到攻击密码方案的目的。通过采取系统矩阵的形式, 改进后的私钥为 (V, W^{-1}, G_1, A, P) 。本文采用 Polar 码作为方案的底层编码, 其置换矩阵 P 等价类码族较多, 攻击者通过穷举攻击难以在多项式时间内找到正确的 P , 对密文难以进行操作, $cP^{-1}A^{-1} = mSG_1 \oplus eP^{-1}A^{-1} = (c'_1, c'_2, \dots, c'_{n+w})$, 无法恢复出正确的明文。此外, 根据 Liu 方案选择的参数, 攻击者通过穷举的方法找到其他 4 类密钥也是不可行的。所以, 穷举攻击不会对本文方案的安全性造成影响。

(2) 密钥恢复攻击

密钥恢复攻击的基本思想是: 从公钥 G^{pub} 恢复出正确的私钥。密钥恢复攻击是一种结构攻击, 通常是针对特定的编码, 如 QC-LDPC 码、GRS 码等。本文提出的方案插入了随机列, 同时采用 2×2 随机非奇异矩阵 A 对插入的随机列进行混合, 私钥的结构被打乱。混合之后又随机从长度为 $n+w$ 密文中删除 w 列冻结比特列, 由于攻击者无法知道 Polar 码的极化性质, 进一步增加了私钥结构的复杂度, 攻击者在多项式时间内通过公钥难以恢复出正确的私钥结构。Bardet 等^[24]针对 Polar 码提出一种确定最小权重的攻击, 解决了极化码相对于递减单项式码的等价问题。本文方案插入的随机列扰乱了生成矩阵的原有编码结构,

所以文献[24]不会影响本文方案的安全性。因为 Polar 码的结构与汉明循环码、GRS 码和 RM 码结构不同, 针对底层编码为循环码、GRS 码和 RM 码的结构化攻击, 不会威胁本文方案的安全性。

(3) 信息集译码攻击

信息集译码攻击是对 McEliece 编码方案最有效的攻击, 在所有已知的攻击中, 基于信息集译码的攻击拥有最低的计算复杂度。Stern^[25]首先提出对 McEliece 方案的信息集译码攻击, 此后, 攻击者为了提高对编码方案攻击的有效性, 提出许多改进的高效率的信息集译码攻击^[26]。信息集译码攻击不是利用编码的底层结构对密码方案进行攻击的, 而是利用搜索信息集的方法达到攻击的目的。此外, 信息集译码攻击的工作因子随错误向量的增加而增加。对于 RLCE 方案, 信息集译码攻击寻找的是公钥 G^{pub} 的列数, 而不是私钥 G_1 的列数。从含有 t 个错误的 $n+w$ 位密文中随机选择 k 位, 构成 x_k , 从错误向量 e 中选择对应位置的 k 位构成 e_k , 从 G^{pub} 选取对应的列构成 $k \times k$ 矩阵 G'_k 。如果选取的 e_k 不含有错误的比特, 即 $e_k = 0$, 攻击者很容易恢复出明文 m 。

对于从 RLCE 方案公钥中随机选择的 k 列, 密文在这些位置上不包含错误的概率为

$$\frac{C_k^{n+w-t}}{C_k^{n+w}}$$

本文通过插入 w 个随机列, 增加了寻找选取的错误向量 e_k 无差错位置的困难性。给定适当的参数, 攻击者通过信息集译码攻击获取明文 m 是极其困难的, 所以, 本文通过插入随机列保证了密码方案不受信息集译码攻击的影响。

(4) 反应攻击

反应攻击是攻击者通过修改少量的密文 c , 将修改后的密文发送给接收方, 观察接收方是否可以正确译码。对于给定的密文 c , 攻击者随机选择位置 i , 在位置 i 添加错误, 将添加错误的密文 c 发送给接收方, 观测接收方是否可以正确译码。如果接收方可以正确译码, 说明攻击者在位置 i 添加了错误; 假如接收方译码失败, 说明攻击者在位置 i 没有添加错误。重复进行上述操作, 直到攻击者获得 k 个无差错位。本文通过插入 w 个随机列, 消息填充, 可以抵抗反应攻击。Liu

提出的基于 Polar 码的 RLCE 方案没有考虑到反应攻击, 本文对可能存在的反应攻击加以分析发现, 不同于底层编码为汉明循环码的密码方案, 通过消息填充的方法, 本文方案可以抵抗可能存在的反应攻击。

5 IND-CCA2 安全性

本文方案采用系统矩阵对明文 m 进行加密, 使用这种方法加密明文, 可能会导致方案 IND-CCA2 的安全性降低。Liu 提出的基于 Polar 码的 RLCE 方案不具有语义安全, 并举例说明了一种相关性攻击的情况: 对同一明文 m 进行两次加密得到两个不同的密文, 可以将两个不同的密文进行对比分析来找出原始的消息。根据上述阐述的消息填充方法, 本文采用 RLCEspad 填充方法, 对每个密文进行普通编码, 使部分消息被编码在错误向量 e 非零项中。发送方在发送明文 m 之前, 通过 3 个随机谕言机模型 H_1, H_2, H_3 将信息打乱, 将部分消息编码在错误向量 e 非零项中, 然后将编码后的消息发送给发送方。攻击者为了获取明文 m , 攻击填充后的密码方案, 困难性等价于基于 Goppa 码的 McEliece 原始方案。假设解码 RLCE 密文是不可区分的, 可以使用与文献[12]中类似的证明来保证 RLCE-RLCEspad 方案可抵御 IND-CCA2 攻击。所以, 本文采用 RLCEspad 方法进行消息填充后的方案具有语义安全, 可以达到 IND-CCA2 安全。

6 性能分析

下面从公钥尺寸与私钥尺寸两个方面分析本文方案。

(1) 公钥尺寸

$$G^{\text{pub}} = SG_1AP = [I|Q]$$

G^{pub} 的尺寸为 $k \times (n+w)$, 通过采用系统矩

阵, 改进后的 G^{pub} 的尺寸为 $k \times (n+w-k)$ 。

(2) 私钥尺寸

私钥 (V, W^{-1}, G_1, A, P) , V 的尺寸为 $(k-u)u$, W^{-1} 的尺寸为 $u \times u$, G_1 的尺寸为 $k \times (n+w)$, A 的尺寸为 $(n+w) \times (n+w)$, P 的尺寸为 $(n+w) \times (n+w)$ 。通过预计算 W^{-1} 的值, 改进后的方案不再存储 $k \times k$ 阶的矩阵 S , 减少了私钥存储空间。

通过表 1, 分析得到:

1) 在相同的比特安全级别, 本文方案的公钥尺寸和 Liu 方案一样, 但本文方案具有 IND-CCA2 安全性, 且本文对部分私钥采用预计算的方法, 可以明显减少私钥存储空间;

2) 在相同的比特安全级别, 本文基于 Polar 码的加密方案公钥尺寸最小, 在 128 bit 安全级别, 本文方案相较于 HermitianRLCE 方案、GRSRLCE 方案和 GoppaMcEliece 方案公钥尺寸分别减少了 4%、46.5%、47.9%。

7 结束语

利用 Polar 码的等价类码族较多、加解密复杂度低的优点, 本文在 RLCE 方案和 PolarRLCE 加密方案的基础上, 利用 Polar 码作为 RLCE 加密方案的底层编码, 通过采用 RLCEspad 消息填充的方法, 将发送的消息扰乱, 提出一种具有 IND-CCA2 安全的基于 Polar 码改进的 RLCE 方案。本文将公钥矩阵转换为系统矩阵, 减少了公钥尺寸, 对部分私钥进行预计算, 采用存储预计算的值得代替存储私钥 S 的值, 减小了私钥存储空间。改进后的方案依然采用 RLCE 方案的结构, 经过安全分析, 可以抵抗目前已知存在的信息集译码等攻击。将 Polar 码作为底层编码, 相较于基于汉明循环码的方案, 本文方案可以抵抗已知存在的结构化等攻击。

Polar 码因其独特的极化性质和较低的译码复杂度等优点, 日益成为研究的热点。将 Polar 码应用到

表 1 基于编码的 RLCE 加密方案公钥尺寸的对比
Table 1 Comparison of public key Size of RLCE schemes based on encoding

信息集译码	公钥尺寸/kB				
	本文	Polar RLCE ^[14]	Hermitian RLCE ^[13]	GRS RLCE ^[11]	Goppa McEliece ^[8]
128 bit 安全级别	98	98	103	183	188
192 bit 安全级别	256	256	198	440	490
256 bit 安全级别	380	380	313	1203	900

基于编码的密码方案中,具有良好的研究前景。在接下来的工作中,可以进一步研究 Polar 码的加密方案,构造更加适合 5G 通信环境下的轻量级加密方案。此外,研究 RLCE 方案的安全性和构造具有 IND-CCA2 安全的 RLCE 密钥封装机制,也是未来研究的重点。利用 RLCE 方案是否可以构造安全高效的签名方案,这个问题值得以后深入研究。

参考文献:

- [1] SHOR P W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer [J]. Siam Review, 1999, 41(2): 303-332.
- [2] 陈宇航, 贾徽徽, 姜丽莹, 等. 基于 Grover 算法的 ECC 扫描式攻击[J]. 信息安全学报, 2016, 16(2): 28-32.
CHEN Y H, JIA H H, JIANG L Y, et al. ECC scanning attack based on Grover algorithm[J]. Netinfo Security, 2016, 16 (2): 28-32.
- [3] GROVER L K. A fast quantum mechanical algorithm for database search[C]//Symposium on the Theory of Computing. 1996: 212-219.
- [4] 睦晗, 吴文玲. 量子对称密码的研究现状与发展趋势[J]. 电子与信息学报, 2020, 42(2): 287-294.
SUI H, WU W L. Research status and development trend of post-quantum symmetric cryptography[J]. Journal of Electronics and Information Technology, 2020, 42(2): 287-294.
- [5] Post-quantum cryptography[M]. 2019.
- [6] POTII A V, KARPENKO A S. Realization of the post-quantum algorithm of the electronic digital signature[J]. Telecommunications and Radio Engineering, 2017, 76(9).
- [7] 王丽萍, 戚艳红. 基于编码的后量子公钥密码研究进展[J]. 信息安全学报, 2019, 4(2): 20-28.
WANG L P, QI Y H. Recent progress of code-based post-quantum public key cryptography[J]. Journal of Information Security, 2019, 4(2): 20-28.
- [8] MCELIECE R J. A public-key cryptosystem based on algebraic coding theory[J]. DSN Progress Report, 1978, 42(44): 114-116.
- [9] NIEDERREITER H. Knapsack-type cryptosystems and algebraic coding theory[J]. Problems of Control and Information Theory, 1986, 15(2): 159-166.
- [10] BARELLI E. On the security of some compact keys for McEliece scheme[J]. arXiv preprint arXiv: 1803.05289, 2018.
- [11] WANG Y. Quantum resistant random linear code based public key encryption scheme RLCE[C]//2016 IEEE International Symposium on Information Theory (ISIT). 2016: 2519-2523.
- [12] WANG Y. Revised quantum resistant public key encryption scheme RLCE and IND-CCA2 security for McEliece schemes[J]. IACR Cryptology ePrint Archive, 2017, 2017: 206.
- [13] MATTHEWS G L, WANG Y. Quantum resistant public key encryption scheme hermitianRLCE[C]//Code-Based Cryptography Workshop. 2019: 1-10.
- [14] LIU J, WANG Y, YI Z, et al. Quantum resistant public key encryption scheme polarRLCE[C]//International Conference on Algebra, Codes and Cryptology. 2019: 114-128.
- [15] ARIKAN E. Channel polarization: a method for constructing capacity-achieving codes[C]//2008 IEEE International Symposium on Information Theory. 2008: 1173-1177.
- [16] 张胜军, 钟州, 金梁, 等. 基于安全极化码的密钥协商方法[J]. 电子与信息学报, 2019, 41(6): 1413-1419.
ZHANG S J, ZHONG Z, JIN L, et al. Secret key agreement based on secure polar code[J]. Journal of Electronics and Information Technology, 2019, 41(6): 1413-1419.
- [17] ZHANG H, LI R, ZHANG G, et al. Communication method using polar code, and wireless device: U.S. patent application 16/391,327[P]. 2019-8-22.
- [18] SHRESTHA S R, KIM Y S. New McEliece cryptosystem based on polar codes as a candidate for post-quantum cryptography[C]//2014 14th International Symposium on Communications and Information Technologies (ISCIT). 2014: 368-372.
- [19] HOOSHMAND R, SHOOSTARI M K, AREF M R. PKC-PC: a variant of the McEliece public key cryptosystem based on Polar codes[J]. arXiv preprint arXiv: 1712.07672, 2017.
- [20] POINTCHEVAL D. Chosen ciphertext security for any one-way cryptosystem[C]//International Workshop on Public Key Cryptography. 2000: 129-146.
- [21] FUJISAKI E, OKAMOTO T. Secure integration of asymmetric and symmetric encryption schemes[J]. Journal of Cryptology, 2013, 26(1): 80-101.
- [22] EATON E W, YAMADA A, KALACH K. Conversion schemes for public key cryptosystems: U.S. patent application 15/852,816[P]. 2018-8-28.
- [23] WANG Y. RLCE Key Encapsulation mechanism (RLCE-KEM) specification[J]. NIST Submission, 2017.
- [24] BARDET M, CHAULET J, DRAGOI V, et al. Cryptanalysis of the McEliece public key cryptosystem based on polar codes[C]//Post-Quantum Cryptography. 2016: 118-143.
- [25] STERN J. A method for finding code words of small weight[C]//International Colloquium on Coding Theory and Applications. 1988: 106-113.
- [26] JOUNDAN I A, NOUH S, AZOUAZI M, et al. A new efficient way based on special stabilizer multiplier permutations to attack the hardness of the minimum weight search problem for large BCH codes[J]. International Journal of Electrical and Computer Engineering, 2019, 9(2): 1232.

[作者简介]



李喆 (1994—), 男, 安徽宿州人, 武警工程大学硕士生, 主要研究方向为量子密码。



韩益亮 (1977—), 男, 甘肃会宁人, 武警工程大学教授、博士生导师, 主要研究方向为信息安全、舆情分析。



李鱼 (1995—), 男, 重庆丰都人, 武警工程大学硕士生, 主要研究方向为格密码。