基于 Polar 码改进的 McEliece 加密方案

李　喆,韩益亮*,李　鱼,吴立强
(武警工程大学密码工程学院,陕西西安 710086)

摘要:量子计算机计算能力的迅猛发展,对当前基于数论的传统密码方案造成严重威胁。本文利用极化码的极化性质,提出基于 Polar 码改进的 McEliece 加密方案,其密钥尺寸比原始密码方案减少约 62%,改进后的连续消除(SC)译码算法译码正确率显著提高,通过证明本文加密方案达到 IND－CCA2 安全级别,可以抵抗目前已知存在的信息集译码等攻击。

关键词:量子计算;Polar 码;McEliece;SC 译码算法; IND－CCA2
中图分类号:TP391.9　　**文献标识码**:A

Improved McEliece encryption scheme based on Polar code

Li Zhe, HAN Yiliang*, Li Yu, Wu Liqiang
College of Password Engineering, Engineering University of PAP, Xi'an Shaanxi 710086

ABSTRACT: The rapid development of quantum computer computing power poses a serious threat to the traditional cryptographic schemes based on number theory. In this paper, using the nature of the polarization of the code, McEliece encryption scheme based on Polar code improvement is put forward, the key size than the original password scheme to reduce about 62%, successive elimination (SC) decoding algorithm can improve the decoding accuracy significantly increased, by showing that level of encryption scheme in this paper to achieve IND－CCA2 proof, can resist known attacks, such as information collection of decoding.

KEYWORDS: Quantum computing; polar codes; McEliece cryptosystem; SC decoding algorithm; IND－CCA2

0　引言

随着量子信息科学的建立,量子计算机的理论和物理实现蓬勃发展,取得许多重要的成果。1994 年,shor[1] 提出了有效分解大整数问题和离散对数问题的量子算法,shor 算法的提出对传统的密码体制造成了极大的安全威胁。虽然量子计算机还没有投入到实际应用中,但是一旦量子计算机真正投入使用,会对现行的密码方案造成极大的威胁,我们应该未雨绸缪,提前做好量子时代到来的准备,需要迫切寻找可以抵御量子计算机的新型密码体制,称之为抗量子计算密码体制(Post Quantum Cryptography)。能抵抗量子计算机攻击的密码体制主要有四种[2]:(1)基于编码的公钥密码体制、(2)基于 hash 的公钥密码体制、(3)基于格的公钥密码体制、(4)基于多变量的公钥密码体制。2015 年、2016 年,NSA 和 NIST 分别推动抗量子算法的标准化[3],2018 年,NIST 举办征求抗量子算法的国际会议,在候选的密码体制中,大约 3/8 的密码体制都是基于编码的密码体制[4],由此可见,基于编码的密码体制是很有研究价值的方案之一。

为了减小 McEliece 体制的公钥长度,许多学者相继提出了许多关于 McEliece 体制的变型,基本的思想是寻找具有更加紧凑的生成矩阵或校验矩阵的码字代替原始方案的 Goppa 码[5]。Polar 码是目前为止唯一可以在理论上证明无限接近于香农限的编码。2009 年,土耳其 Arikan[6] 教授提出 Polar 码后,近年来成为研究的热点,越来越多的学者开始研究 Polar 码的结构及性能。迄今为止,利用极性码的性质,人们在信息安全理论方面进行了许多尝试[7-8]。R. Hooshmand 等[9] 通过对有限长度的极化码性能进行分析,提出了基于物理层加密(PLE)方案。在 2014 年,S. R. Shrestha 等[10] 提出了基于 Polar 码的 McEliece 密码方案。R. Hooshmand 等[11] 提出了对基于 Polar 码的 McEliece 密码方案的进一步优化,减少了密钥长度。但是,在 2016 年,Bardet[12] 等针对文献[10]中提出的基于 Polar 码的 McEliece 密码方案进行结构分析,提出了密钥恢复攻击的方法。这种攻击方法可以获得 Shrestha－Kim 方案解密密文所需要的任何信息。抵御这种密钥恢复攻击的唯一方法是找到极化码的参数,对于极化码参数,找到最小重量码字是不可能的。我们需要对文献[10]提出的参数进行进一步优化选择,以获得更加安全可靠的密码体制。

基金项目:国家自然科学基金资助项目(61572521)

本文针对 McEliece 原始方案存在的不足，提出了一种安全可靠的基于极化码的密码方案。相比 McEliece 原始方案及其他变体，基于极化码的密码方案有以下优点：(1)极化码比 Goppa 码等其他码字具有更好的纠错能力，会大大减少密钥长度；(2)极化码的连续消除(SC)译码算法比 Goppa 码的译码效率明显提高，降低了解密过程中的计算复杂度；(3)极化码具有较大的等价码族，攻击基于极化码的方案来解决码字的等价问题是不可行的。

1 基础知识

Polar 码是根据其信道极化现象构造。信道极化现象主要包括信道联合和信道分裂两部分。通过信道联合和信道分裂后，各个子信道的对称容量将呈现两级分化的趋势：随着码长的增加，出现信道容量趋近于 1 的无噪信道和信道容量趋近于 0 的全噪信道，信道容量趋近于 1 的无噪信道用来传输信息比特，信道容量趋近于 0 的全噪信道用来传输固定比特(冻结比特)。

1.1 相关定义

一个二进制输入离散无记忆信道($B-DMC$)可以表示为 $W:X\to Y$，X 是输入符号集合，Y 是输出符号集合，转移概率为 $W(y\mid x)$，$x\in X$，$y\in Y$。对信道 W 的 N 次极化后的信道可以表示为 W^N，则信道 $W^N:X^N\to Y^N$ 的转移概率为

$$W^N(y_1^N\mid x_1^N)=\prod_{i=1}^{N}W(y\mid x)$$

对于一个 B-DMC，有两个重要的信道容量参数：

(1)对称容量(Symmetric Capacity)：

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

(2)巴氏参数(Bhattacharyya Parameter)：

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

$I(W)$ 是对信道速率的度量，$Z(W)$ 是对信道可靠性的度量。信道 W 在等概率输入情况下，可靠传输时的最大速率为 $I(W)$；而信道 W 在只传输 0 或 1 的情况下，最大似然判决错误概率的上限为 $Z(W)$。

$I(W)$ 与 $Z(W)$ 满足这样的关系：当且仅当 $Z(W)\approx 0$ 时，$I(W)\approx 1$；当且仅当 $Z(W)\approx 1$ 时，$I(W)\approx 0$。

1.2 信道极化

信道极化分为两个阶段：信道联合阶段(Channel Combining)和信道分裂(Channel Splitting)阶段。

1.2.1 信道联合

联合 $B-DMC$ W 的 N 个独立信道，通过递归方式产生一个向量信道 $W^N:X^N\to Y^N$，其中 N 为 2 的次幂 $N=2^n$，≥ 0。$uN\,1\to xN\,1$ 是由复合信道 W_N 的输入到原始信道 W^N 的输入的映射。因此有 $uN\,1=xN\,1G_N$。其中 $uN\,1$ 为原始比特序列，$xN\,1$ 为编码后的比特序列，G_N 为 N 维生成矩阵，码长为 $N=2^n$。

信道 W_N 和 W^N 的转移概率有如下关系：

$$W_N(y_1^N\mid u_1^N)=W^N(y_1^N\mid u_1^NG_N)$$

其中 $yN\,1\in Y^N$，$uN\,1\in X^N$。

1.2.2 信道分裂

将信道联合构成的复合信道 W_N 分裂为 N 个二进制输入的坐标信道(Coordinate Channels)$W(i)$ $N:X\to Y^N\times X^{i-1}$，$1\leq I\leq N$，定义其转移概率为

$$W_N^{(i)}(y_1^N,u_1^{i-1}\mid u_i)\triangleq\sum_{u_{i+1}^N\in X^{N-i}}\frac{1}{2^{N-1}}W_N(y_1^N\mid u_1^N)$$

其中($yN\,1$，$ui-1$ N)表示 $W(i)$ N 的输出，u_i 表示 $W(i)$ N 的输入。

1.3 信道极化定理

对任意 $B-DMC$ 与任意 $\delta\in(0,1)$，当 N 以 2 的幂次趋近于无穷大时，极化信道 $W(i)$ N 中，满足 $I(W_N^{(i)})\in(1-\delta,1)$ 的信道数占总信道数 N 的比例趋于 $I(W)$；满足 $I(W_N^{(i)})\in(0,\delta)$ 的信道所占的比例趋于 $1-I(W)$。

1.4 极化码编码原理

极化编码的基本思想是：只在 $Z(W(i)$ $N)$ 趋近于 0($I(W(i)$ $N)$)趋近于 1)的坐标信道 $W(i)$ N 上发送数据比特。

极化编码步骤：

(1) 极化信道可靠性估计

(2) 比特混合

(3) 构造生成矩阵

1.4.1 极化信道可靠性估计

$Z(W_N^{(i)})$

$$=\sum_{y_1^N\in y^N}\sum_{u_1^{i-1}\in X^{i-1}}\sqrt{W_N^{(i)}(y_1^N,u_1^{i-1}\mid 0)W_N^{(i)}(y_1^N,u_1^{i-1}\mid 1)}$$

$Z(W(i)$ $N)$ 越小，相对应的对称容量 $I(W(i)$ $N)$ 越大，说明信道比较可靠，是好信道，适合传输信息比特；反之，$Z(W(i)$ $N)$ 越大，相对应的对称容量 $I(W(i)$ $N)$ 越小，说明信道不可靠，是坏信道，适合传输固定比特(冻结比特)。

1.4.2 比特混合

消息比特由相对可靠性高的 K 个分裂子信道进行传输，而冻结比特由其他不太可靠的其他分裂子信道传输冻结比特。这一步的输出为编码的原始比特 $uN\,1$。对于 BEC 来说，选择巴氏参数 $Z(W(i)$ $N)$ 最小的 K 个子信道放置消息比特。

1.4.3 构造生成矩阵

极化码通过构造一种编码体制，只选择其中 $Z(W(i)$ $N)$ 接近于 0 的那些信道来发送信息比特。首先声明定义 Kronecker 矩阵乘法：

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

定义 F 矩阵为 $F=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}$，则有生成矩阵 $G_N=B_NF^{\otimes n}$，其中 $F^{\otimes n}$ 表示对矩阵 F 的 n 次克罗内克积，$F^{\otimes n}=F\otimes$

$F^{\otimes(n-1)}$。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

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

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序元素和偶序元素，即先对奇序元素排列，再对偶序元素排列，例如：

$$(u_1, u_2, u_3, u_4 \ldots u_N) \times R_N = (u_1, u_3, u_5, u_{N-1}, u_2, u_4, u_6, u_N)$$

2 编码方案

2.1 McEliece 加密体制介绍

1978 年，McEliece 首先提出一种基于纠错码的密码体制，该体制采用二元 Goppa 码作为原始方案。McEliece 体制的安全性是基于一般线性码译码的困难问题，这个困难问题可以归约到 NPC 问题。一直到现在，McEliece 体制的原始方案依然安全。与 RSA 等相比，该密码体制仅涉及矩阵运算，加解密速度较快，并且可以抵御量子计算机的攻击，但是该体制密钥量大，码率低，因此并未投入到实际使用。

系统参数：n ，$t \in N$，其中 $t \ll n$。

(1)密钥生成过程：对于给定的参数 n 和 t，产生下列矩阵。

G：域 F 上的信息为数为 k、最小距离 $d \geq 2t+1$ 的码 C 的 $k \times n$ 阶生成矩阵，并且可以纠正 t 个错误。

S：$k \times k$ 阶二元随机非奇异可逆矩阵。

P：$n \times n$ 阶二元随机置换矩阵。

然后计算 $k \times n$ 阶矩阵 $G^{pub} = SGP$

公钥：(G^{pub}, t)

私钥：(S, D_C, P)，其中 D_C 是码 C 的一个有效译码算法。

(2)加密过程：$(E_{(Gpub,t)})$

要加密一个明文 $m \in F^k$，选择一个重量为的随机向量 $e \in F^n$，e 为重量小于等于 t 的随机差错图样，如下计算密文：

$$c = mG^{pub} \oplus e$$

(3)解密过程：$(D_{(S,DC,P)})$

要解密一个密文 c，首先计算：

$$cP^{-1} = (mS)G \oplus eP^{-1}$$

然后用对其进行译码，因为可以看成码的一个含有 t 个错误的码字，所以可以译码得到

$$mSG = D_C(cP^{-1})$$

求解等式 $m_0 G = mSG$，得到 m_0，再计算 $m = m_0 S^{(-1)}$。

2.2 Polar 加密

对于参数为 ( )的 *Polar* 码，其中 N 为码长，K 为信息比特 u_A 长度，$N-K$ 为冻结比特长度。

$u = (u_A, u_{AC})$

$x = u_A G_A + u_{AC} G_{AC} = u_A G_A + c$

$c \triangleq u_{AC} G_{AC}$

$R = u_A / x = K/N$

2.3 SC 译码算法

K 个信息比特 u_A 和 $N-K$ 个冻结比特 u_{AC} 编码后经过信道 W^N 进行传输，在译码端接收到传输后的序列 $yN\,1$，且传输概率为 $W_N(yN\,1 | uN\,1)$。*SC* 译码算法就是通过获取 $yN\,1$ 的值以及冻结比特 u_{AC} 的索引位置，来进行译码判决，获取 $uN\,1$ 的估计值 $\hat{u}_1^N$。其 *SC* 译码器包含 N 个译码单元，每一个单元对应一个比特的判决估计。

首先定义对数似然比(Log - Likelihood Ratio，LLR)

$$L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) \triangleq ln\left(\frac{W_N^{(i)}(y_1^N, u_1^{i-1} | 0)}{W_N^{(i)}(y_1^N, u_1^{i-1} | 1)}\right)$$

因此，进行 *SC* 译码时，对于冻结比特可以直接对其进行判决，

$$\begin{cases} \hat{u}_i = u_i, i \in A^C \\ h_i(y_1^N, \hat{u}_1^{i-1}), i \in A \end{cases}$$

即当 $i \in A^C$ 时，表明该比特为冻结比特，即收发方事先约定好的比特，直接对冻结比特估计值赋值 $i \in A^C$。而当 $i \in A$，表明该比特是信息比特，判决函数为：

$$h_i(y_1^N, \hat{u}_1^{i-1}) = \begin{cases} 0, L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) \geq 0 \\ 1, L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) < 0 \end{cases}$$

i 为奇数时，

$$L_N^{(2i-1)}(y_1^N, \hat{u}^{2i-2}) \triangleq \frac{L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}) \cdot L_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2}) + 1}{L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}) + L_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2})}$$

i 为偶数时，

$$L_N^{2i}(y_1^N, \hat{u}^{2i-1}) \triangleq [L_{N/2}^{(i)}(y_1^{N/2}, \hat{u}_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2})]^{1-2\hat{u}_{2i-1}} L_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2})$$

SC 译码算法步骤如下：

(1)初始化

$yN\,1 = (y_1, y_2, y_3 \cdots y_n)$ 表示接收端得到的序列，对于每一个 y_i 都有

$$L_1^{(1)}(y_i) = \frac{w(y_i | 0)}{w(y_i | 1)}$$

(2)依次计算发送端第 i 个比特的似然值

$\hat{u}_1^N = (\hat{u}_1, \hat{u}_2, \hat{u}_3, \cdots \hat{u}_n)$ 表示接收端对发送端编码前 $u_1^N = (u_1, u_2, u_3, \cdots u_n)$ 的可靠性估值序列

$$L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) \triangleq ln\left(\frac{W_N^{(i)}(y_1^N, u_1^{i-1} | 0)}{W_N^{(i)}(y_1^N, u_1^{i-1} | 1)}\right)$$

(3)进行判决

$$h_i(\hat{u}^i) = \begin{cases} 0, L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) \geq 0 \\ 1, L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) < 0 \end{cases}$$

$$\hat{u}^i = \begin{cases} 0, (y_1^N, \hat{u}_1^{i-1}) \geq 1 \\ 1, (y_1^N, \hat{u}_1^{i-1}) < 1 \end{cases}$$

返回步骤(2)进行下一比特的译码，直到该码字全部译码完毕。

对于任意 $B-DMC$ W，*SC* 译码错误率上限 $P_e \leq \sum_{i \in A} Z(W_N^{(i)})$。在 *SC* 译码算法下，达到可靠通信的条件是 $R < I(W) - n^{-1/u}$，u 为缩放指数。

3 基于 Polar 码改进的密码方案

本文提出的方案和 McEliece 原始方案类似,包括密钥生成(公钥,私钥),加密过程,解密过程三个部分。主要改进是提出了一种有效的子信道分类方法,并隐藏了极化码的生成矩阵。随机从好的信道选择 k 个子信道,称为秘密信息集,用 $A_{(s)}$ 来表示。秘密信息集的生成矩阵是由 $k\times n$ 维矩阵 G_n 中与 $A_{(s)}$ 对应 k 行构成的子矩阵。类似地,秘密冻结信息集是从好信道选择的 $(n-k)$ 的子信道,秘密冻结信息集的生成矩阵是有 $(n-k)\times n$ 维矩阵 G_n 中与对应行构成的子矩阵。使用这种方法,攻击者即使在知道 $Z(W)$,n,k 的情况下,也不能恢复出。

3.1 密钥生成

产生秘密生成矩阵。

生成 $k\times k$ 维非奇异置换矩阵 S。

$n\times n$ 维的置换矩阵 $P=[P^1\mid P^2]$,其中 P^1 是 $n\times k$ 维的子矩阵,P^2 是 $n\times(n-k)$ 的子矩阵。

计算阶矩阵。

私钥:$(,D_C,P)$

公钥:(Q,t)

,I_k 是 $k\times k$ 维的单位矩阵,Q 是 $k\times(n-k)$ 维的子矩阵

3.2 加密过程

要加密一个明文 $m\in F^k$,选择一个重量为的随机向量 $e\in F^n$,e 为重量小于等于 t 的随机差错图样,如下计算密文:

$$c = mG^{pub}\oplus e$$

3.3 解密过程

要解密一个密文 c,首先计算:

$$cP^{-1} = (mS)G_{A(S)}\oplus eP^{-1}$$

然后本文采用改进的译码算法对其进行译码,因为可以看成码的一个含有 t 个错误的码字,所以可以译码得到

$$mSG_{A(S)} = D_C(cP^{-1})$$

求解等式,得到,再计算。

改进的译码算法步骤如下:

(1)集合 C 包含当前节点两条路径 e_i 和 $e*\ i$,分别对 e_i 和 $e*\ i$ 进行初始化,$e_i=0$,$e*\ i=1$。初始化两条总路径 $\hat{u}_1^{i-1}$ 和 $\hat{u}_1^{i-1}$,$\hat{u}_1^{i-1}$ 对应 e_i,$\hat{u}_1^{*i-1}$ 对应 $e*\ i$。

(2)$i<N$ 时,分别计算 $Mi+1$ $N(u_{i+1}=1$ 或 0,$e_i\hat{u}_1^{i-1}$ 或 $e_i^*\hat{u}_1^{*i-1}\mid y_1^N)$,一共有 4 个后验概率值。

(3)比较这 4 个后验概率值,将概率值比较大的两个 u_i 添加到 $\hat{u}_1^{i-1}$ 和 $\hat{u}_1^{*i-1}$ 中,u_{i+1} 分别添加到 e_i 和 $e*\ i$。

(4)遍历 $i<N$,重复(2)(3);当 $i=N$ 时,比较 $\hat{u}_1^{i-1}$,$\hat{u}_1^{*i-1}$,e_i,$e*\ i$ 四个值的后验概率值,选择四个值中最大的后验概率值,作为译码序列。

4 安全性分析

(1)穷举攻击

穷举攻击是一种结构攻击,把可能存在的密钥逐个进行解密,直至密文解密。假如密钥空间足够大,穷举攻击不可能对密码方案造成威胁。在本文提出的密码方案中,由于 k 个子信道是 n 个好信道随机生成的,对于参数为 $(n,\ k)$ 的极化码,等价类码共有。只要参数 $(n,\ k)$ 选择合适的数值,在多项式时间内,攻击者通过穷举的方法找到秘密生成矩阵是不可能的。可逆矩阵和置换矩阵也有很多可能性的选择,在数量上也是极其庞大的。在本方案中,非奇异可逆矩阵的数量。置换矩阵的数量。同样,在多项式时间内,攻击者找到可逆矩阵 S 和置换矩阵 P 也是不可能的。例如参数为的极化码,$N_c\approx 2^{825}$,$N_s\approx 2^{825}$,$N_p\approx 2^{1081}$。

(2)密钥恢复攻击

当密钥恢复攻击作为一种代数攻击进行区分时,需要使用区分器将公钥矩阵和随机二进制矩阵进行区分。2013 年,Faugère 等[12]通过对代数攻击中多项式系统的线性化,构造了一个 Goppa 码区分器,能以很高的码率区分随机码和 Goppa 码。在改进的方案中,攻击者无法从随机生成的公钥矩阵中区分出公钥矩阵,因为公钥矩阵不是极性码的生成矩阵,改进的方案通过可逆矩阵和置换矩阵对原有的生成矩阵进行了隐藏。此外,区分器对本方案中所采用的编码的子域空间并不能进行有效的攻击,因此,改进后的密码方案能抵抗密钥恢复攻击。

(3)信息集译码攻击

ISD 攻击一般是通过在接收端给定码字中的扩展集搜索最小重量的码字来寻找密文中的错误向量。在 ISD 译码攻击的发展过程中,Stern 攻击[13]的效果对密码体制的影响最大,极大地破坏了密码方案的安全性。在本文提出的改进方案中,Stern 攻击的工作因子,其中为算法每执行一次迭代所需要的二进制操作数量,P_s 为成功找到最小重量码字的概率。

$$\mathrm{Cost}_s = \frac{1}{2}(n-k)^2(n+k)+2\binom{k/2}{p}pl+$$

$$2p(n-k)\binom{k/2}{p}^2/2^l$$

$$P_S = \binom{k/2}{p}^2\binom{n-k-l}{w-2p}\Big/\binom{n}{w}$$

$$l = log_2\binom{k/2}{p}$$

其中 $0\leqslant p\leqslant w$,$0\leqslant l\leqslant n-k$。例如参数为的极化码,$(p,\ l)=(5,39)$ 时,$W_H(e)=63$,,表明改进后的密码方案可以达到 140 比特安全级别。只要选择合适的参数,改进后的密码方案能够抵抗信息集译码攻击。

(4)适应性选择密文攻击(CCA2)

在本文提出的改进方案中,,我们采用的系统生成矩阵,在使用这种加密矩阵时,有时会导致对抗适应性选择密文攻击安全性的降低。我们需要保证采用系统生成矩阵在减小密钥尺寸和加解密速度更快的优势下,密码方案的安全性仍

然没有受到影响。关于 McEliece 原始方案,提出了许多把 McEliece 公钥密码体制转为 CCA2 安全的技术[14]。本文主要介绍 Kobara – Imai – conversion[15],这种转换可以在合适的参数选择下,有效的减小数据冗余,使得转换后的密码方案可以达到 McEliece 原始方案的安全性级别。攻击者攻击 Kobara – Imai – 转换后的 CCA2 模型,与攻破 McEliece 原始方案的难度相当,这样确保了改进后的密码方案的安全性。

5 效率分析

5.1 密钥量大小

由于在改进的密码方案中,我们采用的系统生成矩阵代替原来的生成矩阵以此来减小密钥尺寸。我们把(,P)作为私钥进行存储,这是因为,S 都是基于 $A_{(s)}$ 生成的,是 $A_{(s)}$ 的补集,相较于 $A_{(s)}$ 来说,需要更小的存储空间。

我们选择参数为(1024,768)的极化码对改进后的密码方案的公钥量大小和私钥量大小进行计算。

公钥量:

$$M_{pub} = k \times (n-k)$$

对于给定参数的情况下,我们计算公钥量大小约为 24.58 kbytes。

私钥量:

$$M_{\text{sec}} = M_{A\complement_{S)}} + M_p$$
$$M_{A\complement_{S)}} = 10(n-k)$$
$$M_p = n \times (n-k)$$

对于给定参数的情况下,我们计算大小约为 2.56 kbits。

对于给定参数的情况下,我们计算大小约为 32.77 kbits。

对于给定参数的情况下,我们计算私钥量大小约为 35.33 kbits4.42kbytes。

表 1 本文方案与原始方案的对比

scheme	McEliece Cryptosystem	Proposed Scheme
Code	Goppa	Polar
(n,k)	(1024,524)	(1024,768)
	67.07kbytes	24.58kbytes
	102.5kbytes	4.42kbytes

5.2 计算复杂度

计算复杂度主要包括两个部分:加密复杂度和解密复杂度。

5.2.1 加密过程

$$c = mG^{pub} \oplus e$$

加密的复杂性主要取决于矩阵乘法和错误向量。

所以加密复杂度

$$C_{Enc} = C_{mul}(mG^{pub}) + C_{add}(e)$$
$$C_{mul}(mG^{pub}) = O(k \times (n-k))$$
$$C_{add}(e) = O(n)$$
$$C_{Enc} = C_{mul}(mG^{pub}) + C_{add}(e) \approx O(k(n-k))$$

5.2.2 解密过程

(1)本文所采用的延迟译码算法,对于每个节点都采取延时判决,采用后一节点的判决结果来代替当前节点的判决,提高当前节点正确判决的概率。

(2)假设第一个信息比特之前有 L_1 个冻结比特,后面有 L_2 个冻结比特,即 $L_1 + L_2 = n - k$。改进后的延时判决译码算法,需要计算的节点数 $L = 4k + L_1 + 2L_2 - 1$,算法复杂度为 $O(2nlogn - (n-1))$。

(3)改进后的延时译码算法复杂度虽然比原来的 SC 译码算法略有提高,但是相较于原来的译码算法,译码正确率大大提高。

(4)改进后的延时译码算法的复杂度虽然比原来的 SC 译码算法略有提高,但是改进后的译码算法可以接近于最大似然译码,相较于原来的译码算法,译码正确率大大提高。

$$cP^{-1} = (mS)G_{A_{(S)}} \oplus eP^{-1}$$
$$mSG_{A_{(S)}} = D_C(cP^{-1})$$

解密的复杂性主要取决于译码算法和 cP^{-1}。

所以解密复杂度

$$C_{Dec} = C_{mul}(cP^{-1}) + C_{SC}(cP^{-1}) + C_{mul}(u_{A_{(S)}}S^{-1})$$
$$C_{mul}(cP^{-1}) = O(n)$$
$$C_{SC}(cP^{-1}) = O(2nlogn - (n-1))$$
$$C_{mul}(u_{A_{(S)}}S^{-1}) = O(k^2)$$
$$C_{Dec} = C_{mul}(cP^{-1}) + C_{SC}(cP^{-1}) + C_{mul}(u_{A_{(S)}}S^{-1}) \approx O(k^2)$$

6 结论

本文结合编码理论最新进展,提出了一种基于 Polar 码改进的密码编码体制,相较于 McEliece 原始方案,改进后的方案密钥尺寸减少了约 62%。在同等安全比特级别下,采用改进后的译码算法,计算复杂度与原始的 SC 译码算法相比略有提高,但译码错误率大大降低,提高了译码正确率。通过 Kobara – Imai – 转换,改进后的方案可以抵抗扩展性攻击和反应攻击等已知存在的攻击类型。在参数选择恰当的情况下,本文提出的改进方案可以实现安全性和效率的平衡,在保证密码体制安全性的基础上,进一步减小密钥长度,减少密钥存储空间。以后,可以对密码方案参数的选择进行优化,进一步实现安全性与效率的提高,当然本文提出的译码算法效率还有待于进一步提高。

参考文献:

[1] 游伟青,陈小明,齐健. 一类抗量子计算的公钥密码算法研究[J]. 信息网络安全,2017,17(4):53 – 60.

[2] Alagic G, Alagic G, Alperin – Sheriff J, et al. Status report on the first round of the NIST post – quantum cryptography standardization process[M]. US Department of Commerce, National Institute of Standards and Technology, 2019.

[3] NIST Post – quantum cryptography standardization website https://csrc.nist.gov/Projects/Post – Quantum – Cryptography/Round – 1

- Submissions

[4] Bernstein D J, Chou T, Lange T, et al. Classic McEliece: conservative code - based cryptography 30 March 2019[J]. 2019.

[5] E. Arıkan "Channel polarization: A method for constructing capacityachieving codes for symmetric binary - input memoryless channels",IEEE Trans. Inf. Theory, vol. 55, no. 7, pp. 3051 - 3073, 2009.

[6] H. Mahdavifar, A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," IEEE Trans. Inf. Theory, vol. 57, no. 10, pp. 6428 - 6443, 2011.

[7] 杨超, 肖东亮, 顾珍珍, 等. 基于 Polar 码的 Niederreiter 公钥密码体制[J]. 密码学报, 2018, 5(6): 623 - 630.

[8] R. Hooshmand, M. R. Aref, "Polar code - based secure channel coding scheme with small key size," IET Commun., vol. 11, no. 15, pp. 2357 - 2361, 2017.

[9] S. R. Shrestha, Y. S. Kim, "New McEliece cryptosystem based on polar codes as a candidate for post - quantum cryptography" Int. Symp. On Commun. and Inf. Technologies (ISCIT), pp. 368 - 372, 2014.

[10] R. Hooshmand, M. K. Shooshtari, T. Eghlidos, M. R. Aref, "Reducing the key length of McEliece cryptosystem using polar codes", Int. ISC Conf. on Inf. Security and Cryptology (ISCISC), pp. 104 - 108, 2014.

[11] M. Bardet, J. Chaulet, V. Dragoi, A. Otmani, J. - P. Tillich, "Cryptanalysis of the McEliece public key cryptosystem based on polar codes," Int. Workshop on Post - Quantum Cryptography, LNCS, vol. 9606, pp. 118 - 143, 2016.

[12] H. A. Goli, S. H. Hassani, R. Urbanke, "Universal Bounds on the Scaling Behavior of Polar codes", IEEE Int. Symp. Inf. Theory (ISIT), pp. 1957 - 888, 2012.

[13] Faugere JC, Gauthier - Umana V, Otmani A, Perret L, Tillich JP, "A distinguisher for high - rate McEliece cryptosystems", IEEE Trans on Inf. Theory, vol. 59, no. 10, pp. 6830 - 6844.

[14] J. Stern, "A method for finding codewords of small weight", Coding Theory and Applications, pp. 106 - 113, 1989.

[15] N. Döttling, R. Dowsley, J. Müller - Quade, A. C. A Nascimento, "A CCA2 Secure Variant of the McEliece Cryptosystem", IEEE Trans. Inf. Theory, vol. 58, no. 10, pp. 6672 - 6680, 2012.

[作者简介]

李　喆(1994 -),男,安徽宿州人,硕士,主要研究方向:抗量子密码。

韩益亮(1977 -),男,甘肃会宁人,教授,博士,博士生导师,主要研究方向:信息安全,密码学。

李　鱼(1995—),男,重庆丰都人,硕士,主要研究方向:抗量子密码。

吴立强(1987 -),男,陕西蓝田人,讲师,主要研究方向:信息安全,格的研究。