Criptografia Pós-Quântica baseada em Códigos Corretores de Erros

Aluno: Gervásio Protásio dos Santos Neto
Orientador: Prof^{o} Dr. Routo Terada
IME-USP, São Paulo - 2016
gervasio@ime.usp.br

1. Introdução

Atualmente, os padrões de encriptação mais populares no mundo são o RSA e criptossistemas baseados em curvas elípticas (CCEs). Sua segurança baseia-se na suposta dificuldade do **Problema do Logarítmo Discreto** (PLD).

Contudo, em 1994, Peter Shor propôs um algoritmo quântico que era capaz de resolver o PLD em tempo polinomial [6]. Se computadores quânticos se tornarem viáveis e comercialmente disponíveis, esse algoritmo pode ser utilizado para facilmente descobrir as chaves secretas RSA ou de CCEs.

A comunidade de segurança começou então linhas de pesquisa em algoritmos criptográficos de chave pública que não dependem do PLD. Essa área ficou conhecida como *criptografia pós-quântica*.

Dentre as possíveis alternativa pós-quânticas ao RSA e as ECCs, uma das mais promissoras é a que faz uso de *códigos corretores de erros* [1]. O primeiro criptossistema baseado em códigos foi proposto por McEliece [2] e leva seu nome.

Um dos maiores empecilhos para a adoção do criptossistema de McEliece é o tamanho proibitivamente grande das chaves. Neste trabalho estudamos e comparamos as propostas feitas em [3] para obter chaves mais compactas para o Critpossistema de McEliece.

2. Códigos Lineares

Códigos corretores de erros são estruturas algorítmicas que permitem expressar informação (normalmente, sequências de bits) de tal forma que eventuais erros introduzidos possam ser detectados e corrigidos baseados na informação correta restante.

Definição 2.1. *Seja K um corpo finito. Dizemos que um código $C \subset K^n$ é um **código linear** se C é um subespaço vetorial de K^n.*

Definição 2.2. *Se $C \subset K^n$ é um código linear de dimensão k. Uma matriz $k \times n$ G tal que, $\forall m \in K^k$, $mG \in C$ é uma **matriz geradora** de C. Uma matriz $(n-k)\times n$ H tal que $\forall m \in K^n$, $Hm^T = 0$ é uma **matriz de teste de paridade** de C.*

3. Criptossistema de McEliece

O criptossistema de McEliece tem as seguintes chaves pública e privada, respectivamente:

$$K_{\text{pub}} = (SGP, \delta) \quad \text{e} \quad K_{\text{priv}} = (G, \psi, S, P)$$

Onde temos:

- **G**: uma matriz k x n, geradora de um código linear de dimensão k. A família de códigos a qual o código gerado por G pertence pode ter grande impacto sobre a segurança do sistema e o tamanho das chaves obtidas.
- ψ: Uma estrutura capaz de corrigir δ erros do código gerado por G.
- **S**: Uma matriz k x k, inversível
- **P**: Uma matriz de permutação n x n

O algoritmo para criptografia é:

Data: $m \in K^k$, a mensagem a ser criptografada
Result: $c \in K^n$, a mensagem criptografada
Calcule $m' = m\overline{G}$, a palavra do código correspondente a m
Selecione um vetor aleatório $e \in K^n$ de peso δ (o erro)
Calcule $c = m' \oplus e$
return *c*

O algoritmo de decriptografia é:

Data: $c = m\overline{G} \oplus e$, a mensagem criptografada
Result: m, a mensagem original
Calcule $\overline{c} = cP^{-1} = mSG + eP^{-1}$
Use o corretor de erros ψ para remover os erros e obter
Resolva o sistema linear sobredeterminado dado por $mSG = \overline{m}$, obtendo m
return *m*

4. Códigos de Goppa

Definição 4.1. *Seja K um corpo finito e F uma extensão de K. Tomemos $\varphi(x) \in F[x]$ e $L = \{\alpha_0, \ldots, \alpha_{n-1}\} \subset F$, com $\alpha_i \neq \alpha_j$ se $i \neq j$ e $\varphi(\alpha_k) \neq 0$. Podemos definir o seguinte espaço vetorial sobre K:*

$$\Gamma_K(L, \varphi) = \left\{ c \in K^n : \sum_{k=0}^{n-1} c_k (\varphi(\alpha_k))^{-1} \cdot \frac{\varphi(x) - \varphi(\alpha_k)}{x - \alpha_k} = 0 \right\}$$

*$\Gamma_K(L, \varphi)$ é chamado de o **Código de Goppa** sobre K com suporte L e polinômio φ.*

Definição 4.2. *Um código de Goppa $\Gamma_K(L, \varphi)$ é dito **binário irredutível** se $K = \mathbb{F}_2$ (o corpo de dois elementos), $L \subset \mathbb{F}_{2^m}$ e φ é um polinômio irredutível sobre $F_{2^m}[x]$.*

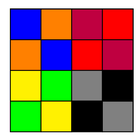
Teorema 4.1. *Um código de Goppa binário irredutível consegue corrigir $grau(\varphi)$ erros.*

Códigos de Goppa binários irredutíveis (**CGBI**) possuem um algoritmo eficiente para correção de erros e formam a classe de códigos utilizada pelo criptossistema de McEliece em sua descrição clássica. Entretanto, as chaves obtidas desta forma são extremamente longas.

5. Códigos de Goppa Quase-Diádicos

Definição 5.1. *Dado um corpo K e um vetor $h = (h_1, \ldots, h_n) \in K^n$, a **matriz diádica** $\Delta(h)$ é a matriz simétrica com componentes $\Delta_{ij} = h_{i \oplus j}$.*

Definição 5.2. *Uma matriz **quase-diádica** é uma matriz (potencialmente não diádica) de blocos, cujos blocos componentes são matrizes diádicas.*



Matriz quase diádica. Cada bloco 4x4 é uma matriz diádica, bem como cada quadrado colorido.

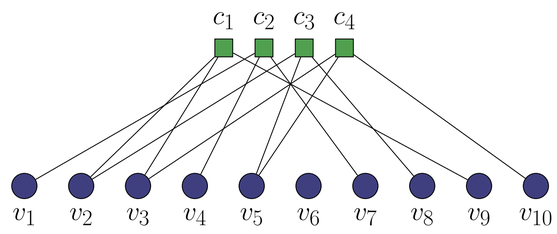
É possível construir códigos de Goppa com poder de correção de erro equivalente aos CGBIs e que admitem uma matriz de teste de paridade quase-diádica. Chama-se essa subclasse de Códigos de Goppa Quase-Diádicos (QD-Goppa)
As simetrias presentes em matrizes quase-diádicas podem ser exploradas para produzir chaves menores que as obtidas com CGBIs clássicos. Além disso, a variente do McEliece que utiliza essa classe de códigos permite tempos de criptografia e decriptografia menores. [4].

6. Códigos QC-MDPC

Códigos MDPC (*moderate density parity check*) são uma família de códigos cuja matriz de teste de paridade é pouco densa (sem ser totalmente esparsa). Esses códigos são desprovidos de estrutura algébrica e podem ser interpretados como grafos bipartidos. Por exemplo, um código MDPC tem matriz de teste de paridade H, e:

$$H = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Então o grafo correspondente, chamado de *Grafo de Tanner* será:



Grafo de Tanner para matriz H: vértices de checagem em verde e de variáveis em azul.

Definição 6.1. *Um código MDPC C de dimensão k sobre $\mathbb{F}_2^n$ é **quase-cíclico** (QC-MDPC) se existe um inteiro η tal que todo shift circular de η bits de uma palavra de C produz outra palavra de C.*

Particularmente, todas as linhas de uma matriz de teste de paridade de C podem ser obtidas por shifts de η bits da primeira

7. Comparação

As variantes que fora estudadas permitem uma grande redução do tamanho da chave. Uma comparação dos tamanhos (em bits) é feita na tabela abaixo, retirada de [5]:

Nível de Segurança	CGBI	QD-Goppa	QC-MDPC
80	460647	20480	4801
128	1537536	32768	9857
256	7667855	65536	32771

Referências

[1] D. Augot, L. Batina, D. J. Bernstein, J. Bos, J. Buchmann, W. Castryck, O. Dunkelman, T. Güneysu, S. Gueron, A. Hülsing, et al. Initial recommendations of long-term secure post-quantum systems. *Available at pqcrypto.eu.org/docs/initial-recommendations.pdf*, 2015.

[2] R. J. McEliece. A public-key cryptosystem based on algebraic coding theory. *Deep Space Network Progress Report*, 44:114–116, 1978.

[3] R. Misoczki. *Two Approaches for Achieving Efficient Code-Based Cryptosystems*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2013.

[4] R. Misoczki and P. S. Barreto. Compact mceliece keys from goppa codes. In *Selected Areas in Cryptography*, pages 376–392. Springer, 2009.

[5] R. Misoczki, J.-P. Tillich, N. Sendrier, and P. S. Barreto. Mdpc-mceliece: New mceliece variants from moderate density parity-check codes. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pages 2069–2073. IEEE, 2013.

[6] P. W. Shor. Polynomial time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1481–1509, 1997.