Tópicos Avançados em Ciência da Computação I: Introdução à Teoria de Códigos para Criptografia Pós-quântica

Daniel Panario
School of Mathematics and Statistics
Carleton University

IC - Unicamp, Sala 351 do IC-3
das 13:30 as 18:30 em 16-17 de janeiro de 2020,
das 13:30 as 17:30 em 20-24 de janeiro de 2020

Introdução à Teoria de Códigos: Parte III

Daniel Panario
School of Mathematics and Statistics
Carleton University

16-24 de janeiro de 2020

Conteúdo da aula

- Códigos cíclicos e polinômios
- Códigos de Hamming
- Códigos quasi-cyclic low density parity check (QC-LDPC)
- Códigos QC-LDPC sem alguns conjuntos de variaveis "trapping"
- Breve comentário sobre códigos MDPC
- Comentários sobre as propostas NIST: LEDAcrypt e BIKE

Códigos cíclicos

Definição

Um código linear $C(n,k)$ sobre $\mathbb{F}_q$ é cíclico, se $(c_0,\ldots,c_{n-1}) \in C$ implica que $(c_{n-1},c_0,\ldots,c_{n-2}) \in C$.

Exemplo: $C = \{000, 110, 101, 011\}$ é um código cíclico.

Códigos cíclicos podem ser caracterizados com polinômios.

Teorema

Um código linear $C(n,k)$ sobre $\mathbb{F}_q$ é cíclico se e somente se C é um ideal de $\mathbb{F}_q[x]/(x^n-1)$.

Retornaremos a esse teorema após ter visto a álgebra de polinômios sobre corpos finitos necessária.

Teorema

Um código linear $C(n,k)$ sobre $\mathbb{F}_q$ é cíclico se e somente se C é um ideal de $\mathbb{F}_q[x]/(x^n-1)$.

Demonstração.

Se C é um código cíclico e $c(x) \in C$, temos que $xc(x), x^2c(x), x^3c(x), \ldots$ também pertencem a C. Seja $a(x) = \sum_i a_i x^i \in \mathbb{F}_q[x]/(x^n-1)$. Como $a(x)c(x) = \sum_i a_i(x^i c(x))$ e C é um subespaço vetorial sobre $\mathbb{F}_q$, temos que C é um ideal.

Reciprocamente, se C é um ideal de $\mathbb{F}_q[x]/(x^n-1)$ e $c(x) = \sum_{i=0}^{n-1} c_i x^i$ é uma palavra-código, então $xc(x)$ também é uma palavra-código. Logo, C é cíclico. □

Teorema

Um código linear $C(n,k)$ sobre $\mathbb{F}_q$ é cíclico se e somente se C é um ideal de $\mathbb{F}_q[x]/(x^n-1)$.

Demonstração.

Se C é um código cíclico e $c(x) \in C$, temos que $xc(x), x^2c(x), x^3c(x), \ldots$ também pertencem a C. Seja $a(x) = \sum_i a_i x^i \in \mathbb{F}_q[x]/(x^n-1)$. Como $a(x)c(x) = \sum_i a_i(x^i c(x))$ e C é um subespaço vetorial sobre $\mathbb{F}_q$, temos que C é um ideal.

Reciprocamente, se C é um ideal de $\mathbb{F}_q[x]/(x^n-1)$ e $c(x) = \sum_{i=0}^{n-1} c_i x^i$ é uma palavra-código, então $xc(x)$ também é uma palavra-código. Logo, C é cíclico. □

Definição

Seja $C = (g)$ um código cíclico. Dizemos que g é o polinômio gerador de C e $h = (x^n - 1)/g$ é o polinômio verificador de C.

Teorema

Seja C um ideal não nulo em $\mathbb{F}_q[x]/(x^n - 1)$, isto é, C é um código cíclico de comprimento n.

1. *O código C é gerado por um único polinômio mônico g de grau mínimo em C.*
2. *O polinômio gerador g de C é um fator de $x^n - 1$.*
3. *Em $\mathbb{F}_q[x]$, qualquer $c \in C$ pode ser escrito unicamente como $c = fg$, onde $\mathrm{grau}(f) < n - r$ e $\mathrm{grau}(g) = r$. Além disso, a dimensão de C é $n - r$. (Assim, a mensagem f se torna a palavra-código fg.)*

Teorema

4. *Se $g(x) = g_0 + g_1 x + \cdots + g_r x^r$, então C é gerado como um subespaço de $\mathbb{F}_q^n$ pelas linhas da matriz geradora*

$$G = \begin{bmatrix} g_0 & g_1 & \cdot & \cdot & \cdot & g_r & 0 & 0 & 0 \\ 0 & g_0 & g_1 & \cdot & \cdot & \cdot & g_r & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & g_0 & g_1 & \cdot & \cdot & \cdot & g_r \end{bmatrix}$$

$$= \begin{bmatrix} g(x) & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & xg(x) & 0 & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 & x^{n-r-1}g(x) \end{bmatrix}.$$

Códigos de Hamming

Definição

Seja m um inteiro maior ou igual a 2. Um código binário C_m de comprimento $n = 2^m - 1$ com uma matriz de paridade H de ordem $m \times (2^m - 1)$ é chamado de código binário de Hamming, se as colunas de H correspondem às representações binárias dos inteiros $1, 2, \ldots, 2^m - 1$.

Exemplo: C_3 tem matriz de paridade

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

A dimensão de C_m é $2^m - 1 - m$. Quaisquer duas colunas são linearmente independentes, já que nenhuma coluna é múltipla da outra. Por outro lado, há três colunas que são linearmente dependentes. Por exemplo,

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 1 \end{pmatrix}.$$

Concluímos que a distância mínima de C_m é $2 + 1 = 3$ e assim C_m corrige um erro.

Teorema

O código de Hamming com parâmetros $n = 2^m - 1$, $k = n - m$ e $d = 3$ é um código com matriz geradora

$$G = \begin{pmatrix} M^{(1)} & 0 & \cdots & 0 \\ 0 & xM^{(1)} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & \cdot & \cdots & x^{n-m-1}M^{(1)} \end{pmatrix},$$

onde o polinômio $M^{(1)} = g$ (o polinômio gerador do código), é o polinômio minimal dos elementos na classe lateral ciclotômica $C_1 = \{1, 2, 4, \ldots\}$ módulo $n = 2^m - 1$.

A seguir veremos um exemplo de código de Hamming construído usando polinômios. Voltaremos aos códigos cíclicos para ver BCH codes e a proposta HQC de NIST após falar de polinômios minimais.

Exemplo

Para $n = 2^3 - 1 = 7$ e $m = 3$, temos que $M^{(1)}(x) = 1 + x + x^3$ é o polinômio minimal dos elementos na classe lateral ciclotômica $C_1 = \{1, 2, 4\}$ módulo 7 e, então, a matriz geradora é

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Se não conhecermos a matriz de paridade H, podemos usar o fato de que $h(x) = (x^7 - 1)/(x^3 + x + 1) = x^4 + x^2 + x + 1$, e, assim, temos que

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

Códigos Quase-Cíclicos (QC) e Códigos QC-LDPC

Códigos quase-cíclicos

Definição

Um código é quase-cíclico (QC) se existe um inteiro s, tal que cada deslocamento cíclico de s posições de uma palavra de código resulta numa palavra de código.

Claramente, um código cíclico é um código quase-cíclico com $s = 1$.

Exemplo: $C = \{0000, 0011, 1100, 0101\}$ é um código quase-cíclico com $s = 2$.

Códigos quase-cíclicos também podem ser caracterizados com polinômios sobre corpos finitos: as palavras de código são polinômios $x^{n-s}c(x)$, para palavras de código $c(x)$ módulo $x^n - 1$ onde n é o comprimento do código.

Exemplo de código quase-cíclico

Exercício: mostrar que a seguinte matriz de paridade é de um código quase-cíclico com $s = 2$:

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

Sugestão.

Primeiro achar todas as palavras do código. Depois, verificar que é um código quase-cíclico com $s = 2$.

O codigo também é quase-cíclico para algum outro valor de s?

Há várias construções desses códigos. Um dos métodos mais comuns nas propostas NIST é baseado no uso de matrizes circulantes.

Códigos quase-cíclicos com matrizes circulantes

Uma matriz circulante é uma matriz quadrada onde cada linha é a linha anterior deslocada uma posição à direita; a última posição torna-se a primeira posição da linha seguinte.

Essas matrizes são completamente definidas pela primeira linha:

$$A = \begin{pmatrix} a_0 & a_1 & \cdots & a_{v-1} \\ a_{v-1} & a_0 & \cdots & a_{v-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}.$$

Uma matriz circulante por blocos é formada pela concatenação de matrizes circulantes quadradas de mesmo tamanho.

Essas matrizes têm peso constante por linha e coluna. Vários métodos dos que estamos interessados neste curso tem peso baixo.

Códigos quase-cíclicos com matrizes circulantes (cont.)

Um dos códigos quase-cíclicos mais simples é formado por várias matrizes circulantes descrevendo a matriz de paridade

$$H = [A_1 \ A_2 \ \cdots A_l],$$

onde $A_1, \ldots, A_l$ são matrizes circulantes binárias $v \times v$.

Se uma das matrizes circulantes é inversível, por exemplo A_l, então a matriz geradora do código pode ser construída em forma sistemática

$$G = \begin{pmatrix} & (A_l^{-1}A_1)^T \\ I_{v(l-1)} & (A_l^{-1}A_2)^T \\ & \vdots \\ & (A_l^{-1}A_{l-1})^T \end{pmatrix},$$

resultando num código quase-cíclico de comprimento vl e dimensão $v(l-1)$.

Códigos quase-cíclicos com matrizes circulantes (cont.)

Observamos que, como uma das matrizes circulantes é inversível, essa construção da matriz geradora necessariamente leva a uma matriz H de posto completo (full rank).

A álgebra de matrizes circulantes binárias $v \times v$ é equivalente à álgebra de polinômios módulo $x^v - 1$ sobre $\mathbb{F}_2$. Uma matriz circulante A é caracterizada pelo polinômio

$$a(x) = a_0 + a_1 x + \cdots + a_{v-1}x^{v-1}$$

com coeficientes da sua primeira linha; então, um código C como o acima é caracterizado completamente pelos polinômios $a_1(x), \ldots, a_l(x)$.

É também necessária a noção de polinômio transposto

$$a(x)^T = \sum_{i=0}^{n-1} a_i x^{n-i}.$$

Códigos quase-cíclicos com matrizes circulantes (cont)

Dado um código binário de comprimento $n = vl$ e dimensão $k = v(l-1)$, os k bits de mensagem $[i_0, i_1, \ldots, i_{k-1}]$ são descritos pelo polinômio

$$i(x) = i_0 + i_1 x + \cdots + i_{k-1} x^{k-1}.$$

A palavra-código para essa mensagem é $c(x) = [i(x), p(x)]$, onde $p(x)$ é dado por

$$p(x) = \sum_{j=1}^{l-1} i_j(x)(a_l^{-1} a_j(x))^T,$$

e $i_j(x)$ é o polinômio representando os bits de informação $i_{v(j-1)}$ to i_{vj-1}, ou seja,

$$i_j(x) = i_{v(j-1)} + i_{v(j-1)+1} x + \cdots + i_{vj-1} x^{v-1}.$$

Observamos que todos os produtos de polinômios são módulo x^{v-1}.

Exemplo: códigos quase-cíclicos com matrizes circulantes

A construção anterior produz códigos QC-LDPC quasi-cyclic low-density parity-check de uso importante em criptografia pós-quântica.

Exemplo: Consideremos um código quase-cíclico de taxa ("rate") $1/2$ com $v = 5$ ($k = 5, n = 10$), construído usando as matrizes circulantes descritas por $a_1(x) = 1 + x$ e $a_2(x) = 1 + x^2 + x^4$. Temos que a matriz de paridade é

$$H = \left(\begin{array}{ccccc|ccccc} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{array}\right).$$

Vemos a presença cíclica dos códigos A_1 e A_2, onde $H = [A_1 \ A_2]$.

Exemplo: códigos quase-cíclicos com matrizes circulantes

Como o segundo polinômio é inversível, a matriz circulante tambem é, e temos um gerador escrito em forma sistemática:

$$a_2^{-1}(x) = x^2 + x^3 + x^4.$$

A matriz geradora contém uma identidade 5×5 (lembre-se que $v = 5$), e temos uma matriz 5×5 descrita pelo polinômio

$$(a_2^{-1} a_1(x))^T = (1 + x^2)^T = 1 + x^3.$$

Então, a matriz geradora é

$$G = \left(\begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{array}\right).$$

Variantes

Um caso importante na prática, usado em alguns sistemas pós-quânticos como BIKE, é aquele em que $l = 2$, ou seja $H = [A_1 \ A_2]$. Nesse caso, o segundo bloco é circulante e simplifica a implementação sem perda de segurança. Se $l > 2$ temos várias partes circulantes na direita da identidade.

Há muitas outras idéias similares. Por exemplo, em LEDAcrypt, uma matriz circulante por blocos é formada pela concatenação de matrizes circulantes quadradas de mesmo tamanho:

$$A = \begin{pmatrix} B_{0,0} & B_{0,1} & \cdots & B_{0,c-1} \\ B_{1,0} & B_{1,1} & \cdots & B_{1,c-1} \\ \vdots & \vdots & \ddots & \vdots \\ B_{r-1,0} & B_{r-1,1} & \cdots & B_{r-1,c-1} \end{pmatrix}.$$

Cada matriz $B_{i,j}$ é uma matriz circulante de permutação.

Exemplo de códigos quase-cíclicos

Exemplo. Consideremos $v = 27$ e um código quase-cíclico de comprimento 108, taxa 3/4, com matriz de paridade:

$$H = [A_1 \ A_2 \ A_3 \ A_4],$$

consistindo de 4 matrizes circulantes de tamanho 27×27 definidas por

$$a_1(x) = 1 + x^3 + x^{16}, \quad a_2(x) = x^2 + x^6 + x^8,$$

$$a_3(x) = x + x^8 + x^9, \quad a_4(x) = x + x^{13} + x^{23}.$$

O polinômio $a_4(x)$ é inversível, com inverso

$$\begin{aligned} a_4^{-1}(x) &= x + x^4 + x^5 + x^6 + x^7 + x^9 + x^{12} + x^{13} \\ &\quad + x^{15} + x^{17} + x^{20} + x^{21} + x^{23} + x^{24} + x^{25}. \end{aligned}$$

Então, podemos escrever a matriz de paridade em forma sistemática:

$$H_S = [A_4^{-1}A_1 \ A_4^{-1}A_2 \ A_4^{-1}A_3 \ I_{27}].$$

Códigos quase-cíclicos com matrizes circulantes (cont)

Temos, então,

$$\begin{aligned} a_4^{-1}(x)a_1(x) &= 1 + x + x^4 + x^5 + x^7 + x^9 + x^{11} + x^{13} + x^{15} \\ &\quad + x^{18} + x^{19} + x^{22} + x^{23} + x^{25} + x^{26}, \\ a_4^{-1}(x)a_2(x) &= x^4 + x^5 + x^7 + x^9 + x^{12} + x^{17} + x^{19} + x^{22} + x^{26}, \\ a_4^{-1}(x)a_3(x) &= x^4 + x^6 + x^7 + x^{10} + x^{13} + x^{14} + x^{15} + x^{18} \\ &\quad + x^{19} + x^{21} + x^{22} + x^{23} + x^{24} + x^{25} + x^{26}, \end{aligned}$$

e a matriz geradora do código é:

$$G = \begin{pmatrix} & (A_I^{-1}A_1)^T \\ I_{81} & (A_I^{-1}A_2)^T \\ & (A_I^{-1}A_3)^T \end{pmatrix}.$$

Observamos que, usando G dessa forma, o código pode ser implementado usando "shift registers"; ver p. 65-66 do texto de Sarah Johnson.

Outros conceitos importantes

Há muitas construções de código quasi-cíclicos. Em geral, se procuram códigos LDPC com cintura de grafo de Tanner pelo menos 6. Mas essa não é a única característica importante do grafo.

Um outro aspecto importante é evitar os conjuntos de vértices chamados de "trapping sets". Mostramos isso com base no artigo QC-LDPC construction free of small size elementary trapping sets based on multiplicative subgroups of a finite field (com F. Amirzade e M. Sadeghi), a ser publicado em Advances in Mathematics of Communications.

Primeiro, brevemente, comentamos outras construções do texto de Sarah Johnson, capítulo 5, p. 71-79, incluindo técnicas baseadas em grafos (algoritmo PEG - Progressive Edge Algorithm), e baseadas em planejamentos combinatórios e geometria finita.

Construction of QC-LDPC Codes Free of Some Elementary Trapping Sets

Propostas NIST: LEDAcrypt e BIKE

Criptografia baseada em códigos

No sistema de McEliece, o tamanho da chave publica é o problema.
Na verdade é um problema...

enorme:

Nível de Segurança	McEliece com código Goppa	RSA
128	1 537 536	3072

Comparação de tamanhos de chave pública (em bits).

A meta principal das propostas ao NIST baseadas em teoria de códigos é reduzir o tamanho das chaves. Para isso, as propostas usam vários tipos de códigos: Low Density Parity Check (LDPC), Moderate Density Parity Check (MDPC), e versões Quase-Cíclicas (QC) desses códigos.

Criptografia baseada em códigos

No sistema de McEliece, o tamanho da chave publica é o problema.
Na verdade é um problema...

enorme:

Nível de Segurança	McEliece com código Goppa	RSA
128	1 537 536	3072

Comparação de tamanhos de chave pública (em bits).

A meta principal das propostas ao NIST baseadas em teoria de códigos é reduzir o tamanho das chaves. Para isso, as propostas usam vários tipos de códigos: Low Density Parity Check (LDPC), Moderate Density Parity Check (MDPC), e versões Quase-Cíclicas (QC) desses códigos.

Criptografia baseada em códigos

No sistema de McEliece, o tamanho da chave publica é o problema.
Na verdade é um problema...

enorme:

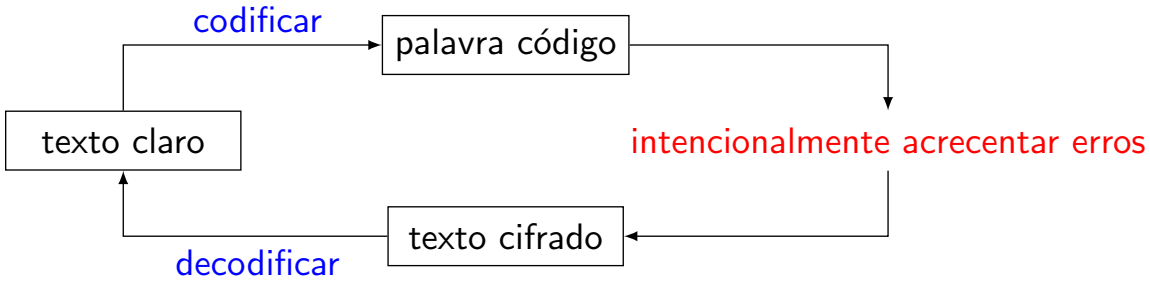
Nível de Segurança	McEliece com código Goppa	RSA
128	1 537 536	3072

Comparação de tamanhos de chave pública (em bits).

A meta principal das propostas ao NIST baseadas em teoria de códigos é reduzir o tamanho das chaves. Para isso, as propostas usam vários tipos de códigos: Low Density Parity Check (LDPC), Moderate Density Parity Check (MDPC), e versões Quase-Cíclicas (QC) desses códigos.

O sistema de McEliece: primeira visão

Robert McEliece (1978) apresentou o primeiro sistema criptográfico baseado na teoria de códigos. A estrutura desse sistema é usada até hoje por outros métodos baseados na teoria de códigos.



O sistema de McEliece: preparação

Elementos do sistema:

- É dado um código de Goppa (binário) Γ de comprimento 1024, dimensão 524 e que pode corrigir 50 erros. O código Γ é mantido em segredo.
- A chave secreta do sistema de McEliece é a tripla (G, S, P) consistindo de uma matriz geradora G para o código Γ, uma matriz 1024×1024 de permutação P, e uma matriz 524×524 inversível S.
- Os tamanhos 1024, 524 e 50 são parâmetros públicos do sistema.
- A chave pública do sistema de McEliece é a matriz 524×1024 $\widehat{G} = SGP$.

O sistema de McEliece: encriptação e decriptação

Encriptação de uma mensagem $m \in \{0,1\}^{524}$:

- Calcular $m\widehat{G}$ e esconder a mensagem somando um vetor de erro aleatório e de comprimento 1024 e peso 50.
- Enviar $y = m\widehat{G} + e$.

Decriptação do texto y:

- Calcular $yP^{-1} = mSG + eP^{-1}$.
- Temos que mSG é uma palavra de código no código secreto Γ; também temos que o vector de erro permutado eP^{-1} tem peso 50.
- Usar o algoritmo de decodificação para o código Γ para achar mS e, portanto, m.

Criptografia baseada em códigos

No sistema de McEliece, o tamanho da chave publica é o problema.
Na verdade é um problema...

enorme:

mas há melhoras...

Nível de Segurança	McEliece com código Goppa	RSA	QC-LDPC	QC-MDPC
128	1 537 536	3072	14 939	11 779

Comparação de tamanhos de chave pública (em bits).

Criptografia baseada em códigos

No sistema de McEliece, o tamanho da chave publica é o problema.
Na verdade é um problema...

enorme:

mas há melhoras...

Nível de Segurança	McEliece com código Goppa	RSA	QC-LDPC	QC-MDPC
128	1 537 536	3072	14 939	11 779

Comparação de tamanhos de chave pública (em bits).

MDPC

Códigos LDPC e MDPC só diferem no peso das linhas: as matrizes MDPC são muito mais densas. Seu peso por linha é $O(n \log n)$.

Essa densidade tão alta leva a uma pior capacidade de correção de erros se certos cuidados não forem tomados.

Por outro lado, essa densidade evita fraquezas causadas pelo uso de matrizes esparsas no sistema; ver:

Using low density parity check codes in the McEliece cryptosystem
de Chris Monico, Joachim Rosenthal a Amin Shokrollahi,
em IEEE International Symposium on Information Theory 2000.

LEDAcrypt e BIKE

Proposta NIST: LEDAcrypt;
seções 1 e 2, páginas 11-26.

Proposta NIST: Bit Flipping Key Encapsulation (BIKE);
seção 1, páginas 1-6.