MDPC-McEliece: New McEliece Variants from Moderate Density Parity-Check Codes

Rafael Misoczki\textsuperscript{1}    Jean-Pierre Tillich\textsuperscript{1}    Nicolas Sendrier\textsuperscript{1}    Paulo Barreto\textsuperscript{2}

\textsuperscript{1}Project SECRET, INRIA-Rocquencourt
              Rocquencourt, France
\textsuperscript{2}Escola Politécnica, University of São Paulo
              São Paulo, Brazil

October 9, 2012 - Dinard, France
Journées Codage et Cryptographie 2012
Outline

LDPC, MDPC and Quasi-Cyclic codes

Previous LDPC McEliece variants

MDPC-McEliece

Security assessment

Benefits

Conclusion
LDPC, MDPC and Quasi-Cyclic codes

Hamming weight

The (Hamming) weight of a vector is the number of its non-zero components.

Binary linear codes

A binary \((n, k)\)-linear code \(C\) of length \(n\) and dimension \(k\) is a \(k\)-dimensional vector subspace of \(\mathbb{F}_2^n\).

- \(C = \langle G \rangle\), where \(G \in \mathbb{F}_2^{k \times n}\) is a generator matrix of \(C\).
- \(C^\perp = \langle H \rangle\), where \(H \in \mathbb{F}_2^{(n-k) \times n}\) is a parity-check matrix of \(C\).

Parity-check matrix density

Let \(w\) be the row weight of a parity-check matrix \(H \in \mathbb{F}_2^{(n-k) \times n}\), its density is the ratio \(w/n\).
LDPC, MDPC and Quasi-Cyclic codes

LDPC and MDPC codes

An \((n, k, w)\)-*DPC code is a linear code which admits a low/moderate density parity-check matrix \(H\) of row weight \(w\).

Decoding

Iterative decoding based on Belief Propagation:
- Error correction capability depends on such density.
- Low complexity for decoding (Bit-Flipping algorithm [Gal63]).

LDPC

- Good trade-off: error correction capability \(\times\) complexity.

MDPC

- Worse error correction capability.
  - Interesting cryptographic features.

There is no known distinguisher for LDPC/MDPC codes.
LDPC, MDPC and Quasi-Cyclic codes

Quasi-cyclic codes

An \((n_0p, k_0p)\)-linear code is quasi-cyclic if any cyclic shift of a codeword by \(n_0\) positions is also a codeword.

We are interested in \(n_0 - k_0 = 1\):

\[
H = \begin{bmatrix} H_0 & H_1 & \cdots & H_{n_0-1} \end{bmatrix},
\]

where \(H_i\) is a \(p \times p\) circulant matrix:

\[
H_i = \begin{bmatrix}
    h_0 & h_1 & h_2 & \cdots & h_{p-1} \\
    h_{p-1} & h_0 & h_1 & \cdots & h_{p-2} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    h_1 & h_2 & h_3 & \cdots & h_0
\end{bmatrix}
\]

When the row weight \(w\) of \(H\) is small: \((n, k, w)\)-QC-*DPC code.

- Compact representation
- Efficient processing (isomorphic to polynomials mod \(x^p - 1\))


McEliece cryptosystem

Let $C$ be an $(n, k)$-linear code able to correct $t$ errors.

Public Key:

$$G \in \mathbb{F}_2^{k \times n}$$

a generator matrix of $C$

Private Key:

$$\Psi$$

t-error correcting procedure for $C$

Encryption:

Let $x \in \mathbb{F}_2^k$, $e \in \mathbb{F}_2^n$, $\text{wt}(e) \leq t$:

$$x \mapsto xG + e$$

Decryption:

Let $y \in \mathbb{F}_2^n$ the received vector:

$$y \mapsto \Psi(y)G^{-1}$$
In [Sen09], a security reduction for Niederreiter cryptosystem. Its security relies on:

**Computational syndrome decoding problem**

Let $t$ be a positive integer.

- Given a matrix $H \in \mathbb{F}_2^{(n-k) \times n}$ and a vector $s \in \mathbb{F}_2^{n-k}$, find a vector $e \in \mathbb{F}_2^n$ of weight $\leq t$ such that $eH^T = s$.

Can be solved with **low weight codeword finding algorithms**.

**Computational low weight codeword finding problem**

- Given an $(n, k)$-linear code $C$ and an integer $t > 0$, find a codeword of $C$ of weight $t$.

Best algorithms: variants of **Information Set Decoding** $^1$.

---

$^1$[Ste89], [BLP08], [FS09], [BLP11], [MMT11], [BJM12]...
McEliece cryptosystem

Its security also relies on:

**Code distinguishing problem**

Let \( \mathcal{F}_{(n,k)} \) be a specific family of \((n, k)\)-linear codes.

- Given a matrix \( G \in \mathbb{F}_2^{k \times n} \), is \( G \) a generator matrix of a code \( C \in \mathcal{F} \)?

- Its hardness depends on the code family.

- Addressed for high rate Goppa codes. [FOPT10]
Previous LDPC/QC-LDPC McEliece variants

LDPC codes ([MRS00])

- **Secret:** an \((n, k, w)\)-LDPC code \(C\).
- **Problem:** Looking for low-weight codewords in \(C^\perp\).

Disguised LDPC codes ([BCG06], [BCGM07], [BC07], [BBC08])

- **Secret:**
  - an \((n, k, w)\)-LDPC code \(C\)
  - a matrix \(Q\) of row weight \(m\)
- **Public-code:** \(C'\) of parity-check \(H' = HQ\).
- **Looking for codewords of weight \(wm\) in \(C'^\perp\) is hard.**
- **Problem:** Constrained structure of \(Q\) weakens [BCG06], [BCGM07], [BC07].
Previous LDPC/QC-LDPC McEliece variants

$Q$ also affects the number of errors:

Private Key: $(H, Q)$

Public Key: $G' = G \cdot Q^{-1}$

$H$: $r \times n$ sparse parity-check matrix of low row weight $w$
$Q$: $n \times n$ sparse circulant matrix of row weight $m$

Encryption:

$y = x \cdot G' + e$
$wt(e) \leq t'$

Decryption:

$y' = y \cdot Q = x \cdot G + e \cdot Q$
Decode $t = mt'$ errors in $y'$. 
Previous LDPC/QC-LDPC McEliece variants

Problems:

LDPC codes:

- Attacks: low weight codeword finding algorithms applied to the dual of the public code.

Disguised LDPC codes:

- Attacks: on the constrained structure of $Q$. 

Misoczki, Tillich, Sendrier, Barreto

MDPC-McEliece

C2 2012
New McEliece Variants from Moderate Density Parity-Check Codes

Solution:

Use MDPC codes (increased weight $w$):

- High enough to avoid low weight codeword attacks on the dual code
- Low enough to allow iterative decoding for a secure amount of errors

Remove the transformation matrices:

- Reduces the venues for mounting structural attacks
New McEliece Variants from MDPC Codes

Key generation

1. Select an \((n, k, w)\)-(QC-)MDPC code of parity-check matrix \(H\)
2. Compute a \(k \times n\) generator matrix \(G\) in systematic form

Private key: \(H\)  
Public key: \(G\)

Encryption

1. Let \(x \in \mathbb{F}_2^k, e \in \mathbb{F}_2^n, \text{wt}(e) \leq t:\)
   \[x \mapsto xG + e\]

Decryption

1. Let \(y \in \mathbb{F}_2^n\) the received vector:
   \[y \mapsto \Psi_H(y)G^{-1}\]
Security assessment

- Security reduction
- Practical security
Security reduction

Decoding problem
- Solved through low weight codeword finding

Distinguishing problem
- Sought structure: sparsity
- Solved through low weight codeword finding

Now, both problems converge to low weight codeword finding!
Practical security

Our proposal: attacks on the dual code of the public code

- The cost of ISD depends on the inverse of the probability of finding a codeword of weight $w$.
- There exist at least $(n - k)$ codewords of weight $w$ on the dual of the public code.

Decoding One Out of Many (DOOM) [Sen11]:

- Attacker possesses multiple instances of the decoding problem and wants to solve only one of them.
Practical security

DOOM:
It gains a factor of $N_s/\sqrt{N_i}$, in comparison with general information set decoding techniques

- $N_i$: Number of available instances of the decoding problem
- $N_s$: Number of solutions of these instances

Example: $N_i = N_s = N$:

$$WF_{doom} = \frac{WF_{isd}}{N_s/\sqrt{N_i}} = \frac{WF_{isd}}{\sqrt{N}}$$
Practical security

Key-distinguishing problem:
- Find one codeword in $C^\perp$ of weight $w$.

Key-distinguishing attacks
- $N_i$: 1 (corresponding to the zero syndrome)
- $N_s$: $r$

MDPC/QC-MDPC case: There is a gain
- Only one low weight codeword is enough to distinguish the code

$$WF_{doom} = \frac{WF_{isd}}{r}$$
Practical security

Key-recovering problem:

- Find $r$ codewords in $C^\perp$ of weight $w$.

Key-recovering attacks

- $N_i$: 1 (corresponding to the zero syndrome)
- $N_s$: $r$

MDPC case: There is no gain

- The attacker must find $r$ low weight codewords

$$WF_{doom} = \frac{WF_{isd}}{r/\sqrt{1}} \cdot r = WF_{isd}$$

QC-MDPC case: There is a gain

- Only one low weight codeword is enough:

$$WF_{doom} = \frac{WF_{isd}}{r}$$
Practical security

Decoding attacks

MDPC case: There is no gain

QC-MDPC case: There is the usual gain of DOOM
  \( N_i = N_s = r \) (all possible cyclic shifts of the syndrome)

\[
WF_{\text{doom}} = \frac{WF_{\text{isd}}}{r/\sqrt{r}} \cdot r = \frac{WF_{\text{isd}}}{\sqrt{r}}
\]
A taste of the QC-MDPC parameters...

<table>
<thead>
<tr>
<th>Security</th>
<th>$n$</th>
<th>$k$</th>
<th>$w$</th>
<th>$t$</th>
<th>pub. key</th>
<th>synd.</th>
<th>dec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>9600</td>
<td>4800</td>
<td>90</td>
<td>84</td>
<td>4800</td>
<td>4800</td>
<td>20ms</td>
</tr>
<tr>
<td>128</td>
<td>19712</td>
<td>9856</td>
<td>142</td>
<td>134</td>
<td>9856</td>
<td>9856</td>
<td>110ms</td>
</tr>
<tr>
<td>256</td>
<td>65536</td>
<td>32768</td>
<td>274</td>
<td>264</td>
<td>32768</td>
<td>32768</td>
<td>1800ms</td>
</tr>
</tbody>
</table>

Public key and syndrome sizes in bits.

Decryption time obtained from a non-optimized C++ implementation running @Intel Xeon CPU @3.20GHz.
Benefits

Security reduction converges to only one (well studied) problem:
  - Low weight codeword finding

Removing the transformation matrices:
  - Reduce the private key size
  - Improve the efficiency of decryption step

QC-MDPC variant:
  - Very compact public-keys

MDPC variant:
  - Further reduces the ways for structural attacks
Conclusion

MDPC codes seem to be very useful for cryptography purposes:
- Less structured than Goppa codes
- Quite close to random linear codes
- Quasi-cyclicity can be successfully applied in order to obtain very small public keys
- Increased density:
  - It avoids attacks on the dual of the public code
  - It approximates the distinguishing problem to the low weight codeword finding problem.

Future works:
- Random linear codes in public-key cryptography!
- Implementation issues
- ...
Questions?

Thanks for your attention!

rafael.misoczki@inria.fr
References


**Misoczki, Tillich, Sendrier, Barreto**