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

Rafael Misoczki[1] Jean-Pierre Tillich[1]
Nicolas Sendrier[1] Paulo Barreto[2]

[1]Project SECRET, INRIA-Rocquencourt
Rocquencourt, France
[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 $\mathcal{C}$ of length n and dimension k is a k-dimensional vector subspace of $\mathbb{F}_2^n$.

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

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 X 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 $n_0 - k_0 = 1$: $H = [H_0|H_1|\dots|H_{n_0-1}]$, where H_i is a $p \times p$ circulant matrix:

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

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 $\mathcal{C}$ 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 $\mathcal{C}$

Private Key:

$$\Psi$$

t*-error correcting procedure for* $\mathcal{C}$

Encryption:

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

$$x \mapsto xG + e$$

Decryption:

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

$$y \mapsto \Psi(y)G^{-1}$$

McEliece cryptosystem

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 $\mathcal{C}$ and an integer $t > 0$, find a codeword of $\mathcal{C}$ 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 $\mathcal{C} \in \mathcal{F}$?

- 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 $\mathcal{C}$.
- Problem: Looking for low-weight codewords in $\mathcal{C}^\perp$.

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

- Secret:
  - an (n, k, w)-LDPC code $\mathcal{C}$
  - a matrix Q of row weight m
- Public-code: $\mathcal{C}'$ of parity-check $H' = HQ$.
- Looking for codewords of weight wm in $\mathcal{C}'^\perp$ 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.

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* $x \in \mathbb{F}_2^k$, $e \in \mathbb{F}_2^n$, $wt(e) \leq t$:
   $x \mapsto xG + e$

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 $\mathcal{C}^\perp$ 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 $\mathcal{C}^\perp$ 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)

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

A taste of the QC-MDPC parameters...

Security	n	k	w	t	pub. key	synd.	dec.
80	9600	4800	90	84	4800	4800	20ms
128	19712	9856	142	134	9856	9856	110ms
256	65536	32768	274	264	32768	32768	1800ms

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

BBC08 M. Baldi, M. Bodrato, and F. Chiaraluce. A new analysis of the mceliece cryptosystem based on qc-ldpc codes. In Proceedings of the 6th interna- tional conference on Security and Cryptography for Networks, SCN 08, pages 246262. Springer-Verlag, Berlin, Heidelberg, 2008. ISBN 978-3-540- 85854-6. doi:http://dx.doi.org/10.1007/978-3-540-85855-3 17.

BC07 M. Baldi and F. Chiaraluce. Cryptanalysis of a new instance of mceliece cryptosystem based on qc-ldpc codes. In Information Theory, 2007. ISIT 2007. IEEE International Symposium on, pages 2591 2595. june 2007. doi:10.1109/ISIT.2007.4557609.

BCG06 M. Baldi, F. Chiaraluce, and R. Garello. On the usage of quasi-cyclic low- density parity-check codes in the mceliece cryptosystem. In Proceedings of the First International Conference on Communication and Electronics (ICEE06), pages 305310. October 2006.

BCGM07 M. Baldi, F. Chiaraluce, R. Garello, and F. Mininni. Quasi-cyclic low- density parity-check codes in the mceliece cryptosystem. In Communica- tions, 2007. ICC 07. IEEE International Conference on, pages 951 956. june 2007. doi:10.1109/ICC.2007.161.

BLP08 D. J. Bernstein, T. Lange, and C. Peters. Attacking and defending the mceliece cryptosystem. In Proceedings of the 2nd International Workshop on Post-Quantum Cryptography, PQCrypto 08, pages 3146. Springer- Verlag, Berlin, Heidelberg, 2008. ISBN 978-3-540-88402-6. doi:10.1007/ 978-3-540-88403-3 3.

BLP11 D. Bernstein, T. Lange, and C. Peters. Smaller decoding exponents: Ball-collision decoding. In P. Rogaway, editor, Advances in Cryptology CRYPTO 2011, volume 6841 of Lecture Notes in Computer Science, pages 743760. Springer Berlin / Heidelberg, 2011. ISBN 978-3-642-22791-2. 10.1007/978-3-642-22792-942.

CC98 A. Canteaut and F. Chabaud. A new algorithm for finding minimum- weight words in a linear code: application to mcelieces cryptosystem and to narrow-sense bch codes of length 511. Information Theory, IEEE Transac- tions on, 44(1):367 378, jan 1998. ISSN 0018-9448. doi:10.1109/18.651067.

Gal63 R.G. Gallager: Low-Density Parity-Check Codes. M.I.T. Press (1963).

FOPT10 J.-C. Faugére, A. Otmani, L. Perret, and J.-P. Tillich. A distinguisher for high rate mceliece cryptosystems. Yet Another Conference on Cryptography (YACC'10), 2010, Porquerolles Island, France.

MRS00 C. Monico, J. Rosenthal, and A. Shokrollahi. Using low density parity check codes in the mceliece cryptosystem. In Information Theory, 2000. Proceedings. IEEE International Symposium on, page 215. 2000. doi:10. 1109/ISIT.2000.866513.

Sen09 N. Sendrier. On the use of structured codes in code based cryptography. In S. Nikova, B. Preneel, and L. Storme, editors, Coding Theory and Cryp- tography III, Contactforum, pages 5968. Koninklijke Vlaamse Academie van Belgi voor Wetenschaeppen en Kunsten, 2009.

Sen11 N. Sendrier. Decoding one out of many. In B.-Y. Yang, editor, Post- Quantum Cryptography, volume 7071 of Lecture Notes in Computer Sci- ence, pages 5167. Springer Berlin / Heidelberg, 2011. ISBN 978-3-642- 25404-8. 10.1007/978-3-642-25405-5-4.

Ste89 J. Stern. A method for finding codewords of small weight. In G. Cohen and J. Wolfmann, editors, Coding Theory and Applications, volume 388 of Lecture Notes in Computer Science, pages 106113. Springer, 1989.