DPA on the 'Secure' Permutation in the McEliece PKC

Tania RICHMOND

Joint work with Martin Petrvalský, Miloš Drutarovský, Pierre-Louis Cayrel and Viktor Fischer

IMATH Laboratory
University of Toulon

RAIM 2016
Banyuls-sur-mer
June 30, 2016

Outline

Outline

Communication

Once upon a time ...

a woman,

Alice

and a man,

Bob

who wanted to communicate together.

Attack

But,
they did not want that anyone,

Eve

could understand this message.

Cryptology

That is why, they use cryptology, i.e.,
the **science of secret**.

kryptos=secret/hidden *logos*=science

Cryptology

Two concepts :

Cryptography

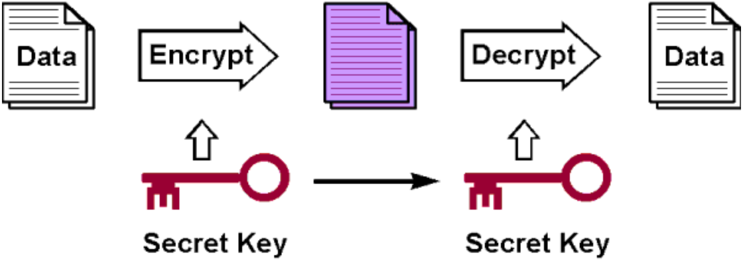
"Secret writting"

Cryptanalysis

"Analysis of a secret message (cryptogram)"

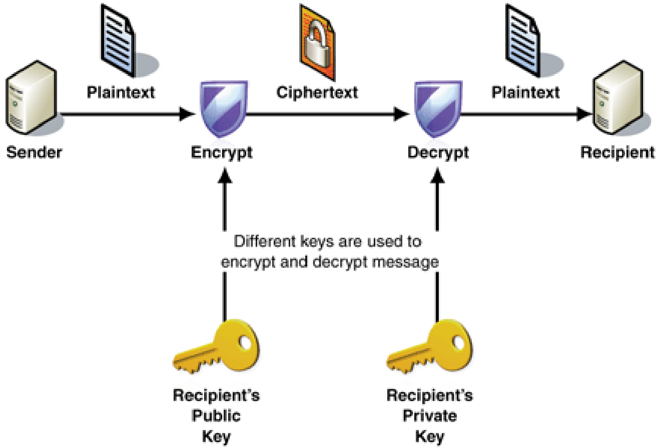
Symmetric Cryptography

Ceasar



Asymmetric Cryptography

[DH76]



McEliece cryptosystem

[McE78]

- First code-based cryptosystem,
- proposed by Robert McEliece in 1978,
- originally using classical Goppa codes.

Linear code

Definition (Linear code)

Let $\mathbb{F}_q$ denoted the finite field of q elements. A linear code $\mathscr{C}$ of length n and dimension k is a k-dimensional subspace of $\mathbb{F}_q^n$.

Definition (Generator matrix)

Let $\mathscr{C}$ be a $[n, k]_q$-linear code. Let $\mathcal{G} \in \mathscr{M}_{k,n}(\mathbb{F}_q)$. We call $\mathcal{G}$ a generator matrix of $\mathscr{C}$ iff $\mathcal{G}$-rows are basis vectors of $\mathscr{C}$.

McEliece key generation

[McE78]

Inputs: n and t two integers.

1. Choose a linear code $\mathscr{C}$ of length n and t-correcting.
   k : dimension of $\mathscr{C}$.
2. Take one generator matrix $\mathcal{G} \in \mathscr{M}_{k,n}(\mathbb{F}_2)$ of $\mathscr{C}$.
3. Randomly choose one invertible matrix $\mathcal{S} \in \mathscr{M}_{k,k}(\mathbb{F}_2)$.
4. Randomly choose one permutation matrix $\mathcal{P} \in \mathscr{M}_{n,n}(\mathbb{F}_2)$.
5. Compute the generator matrix given by $\tilde{\mathcal{G}} = \mathcal{S} \cdot \mathcal{G} \cdot \mathcal{P}$.
6. $s_k \leftarrow (\mathcal{S}, \mathcal{G}, \mathcal{P}, \mathscr{C})$
7. $p_k \leftarrow (\tilde{\mathcal{G}}, t)$
8. Return (p_k, s_k).

Outputs: Public key $p_k = (\tilde{\mathcal{G}}, t)$ and private key $s_k = (\mathcal{S}, \mathcal{G}, \mathcal{P}, \mathscr{C})$.

McEliece encryption

[McE78]

Inputs: Public key $p_k = (\tilde{\mathcal{G}}, t)$, message $M \in \mathbb{F}_2^k$.

1. Encode the message $C = M \cdot \tilde{\mathcal{G}}$.
2. Randomly choose an error vector $E \in \mathbb{F}_2^n$ of weight $w_H(E) = t$.
3. Compute $\tilde{C} = C \oplus E$.
4. Return $\tilde{C}$.

Output: Ciphertext $\tilde{C} \in \mathbb{F}_2^n$ associted to M.

McEliece decryption

[McE78]

Inputs: Private key $s_k = (\mathcal{S}, \mathcal{G}, \mathcal{P}, \Gamma)$, ciphertext $\tilde{C} \in \mathbb{F}_2^n$.

1. Compute $\tilde{C}_p = \tilde{C} \cdot \mathcal{P}^{-1}$.
   i.e. $\tilde{C}_p = M \cdot \mathcal{S} \cdot \mathcal{G} \oplus E \cdot \mathcal{P}^{-1}$
2. Decode $\tilde{C}_p$ to obtain $M \cdot \mathcal{S} \cdot \mathcal{G}$.
3. Get $\tilde{M} = M \cdot \mathcal{S}$ from $M \cdot \mathcal{S} \cdot \mathcal{G}$.
4. Compute $M = \tilde{M} \cdot \mathcal{S}^{-1}$.
5. Return M.

Output: Plaintext $M \in \mathbb{F}_2^k$ associted to $\tilde{C}$.

Side-Channel Attack (SCA)

Differential Power Analysis (DPA)

Definition (SCA)

Exploit the laws of physics phenomenons to obtain some information contained in channels associated to an implementation (software or hardware).

Definition (DPA)

Use several power traces for a same secrete/private key, compute the average to avoid noise (very often), and find a pattern on power traces depending on the secrete/private key in order to recover it.

Outline

Context

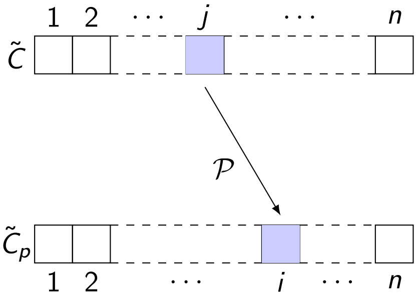
Ciphertext permutation

DPA attack

Conclusion

'Simple' permutation

Example



'Simple' permutation

Algorithm

Inputs: Private permutation matrix $\mathcal{P}^{-1} \in \mathscr{M}_{n,n}(\mathbb{F}_2)$ represented by a lookup table $t^{\mathcal{P}^{-1}}$, ciphertext $\tilde{C} \in \mathbb{F}_2^n$.

For $i = 0$ **to** $n-1$

$$j = t_i^{\mathcal{P}^{-1}}$$
$$\tilde{C}_{p_i} = \tilde{C}_j$$

Endfor

Return $\tilde{C}_p$.

Output: Permuted ciphertext $\tilde{C}_p \in \mathbb{F}_2^n$.

'Secure' permutation [STMOS08]

Algorithm

Inputs: Private permutation matrix $\mathcal{P}^{-1} \in \mathscr{M}_{n,n}(\mathbb{F}_2)$ represented by a lookup table $t^{\mathcal{P}^{-1}}$, ciphertext $\tilde{C} \in \mathbb{F}_2^n$.

For $i = 0$ **to** $n-1$
    $j = t_i^{\mathcal{P}^{-1}}$
    $\tilde{C}_{p_i} = 0$
    For $h = 0$ **to** $n-1$
        $k = \tilde{C}_{p_i}$
        $\mu = \tilde{C}_h$
        $s = j \oplus h$
        $s \mid= s \gg 1$
        $s \mid= s \gg 2$
        $s \mid= s \gg 4$
        $s \mid= s \gg 8$
        $s \mid= s \gg 16$
        $s \mathbin{\&}= 1$
        $s = \sim(s-1)$
        $\tilde{C}_{p_i} = (s \mathbin{\&} k) \mid ((\sim s) \mathbin{\&} \mu)$
    Endfor
Endfor
Return $\tilde{C}_p$

Output: Permuted ciphertext $\tilde{C}_p \in \mathbb{F}_2^n$.

'Secure' permutation [STMOS08]

Examples

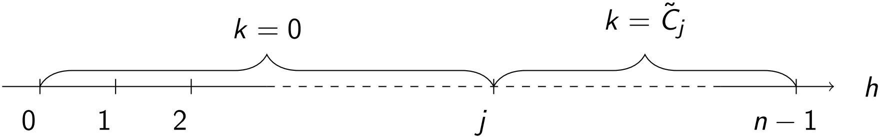
Steps	Test hypotheses			
7: $s = j \oplus h$	$1\underbrace{00\ldots0}_{31}$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{32}$
8: $s \mid= s \gg 1$	$11\underbrace{00\ldots0}_{30}$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{32}$
9: $s \mid= s \gg 2$	$1111\underbrace{00\ldots0}_{28}$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{32}$
10: $s \mid= s \gg 4$	$\underbrace{11\ldots1}_{8}\underbrace{00\ldots0}_{24}$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{32}$
11: $s \mid= s \gg 8$	$\underbrace{11\ldots1}_{16}\underbrace{00\ldots0}_{16}$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{32}$
12: $s \mid= s \gg 16$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{32}$
13: $s \mathbin{\&}= 1$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{00\ldots0}_{31}1$	$\underbrace{00\ldots0}_{32}$
14: $s = \sim(s-1)$	$\underbrace{11\ldots1}_{32}$	$\underbrace{11\ldots1}_{32}$	$\underbrace{11\ldots1}_{32}$	$\underbrace{00\ldots0}_{32}$

Weakness [PRDCF16]

Leakage Step 15:

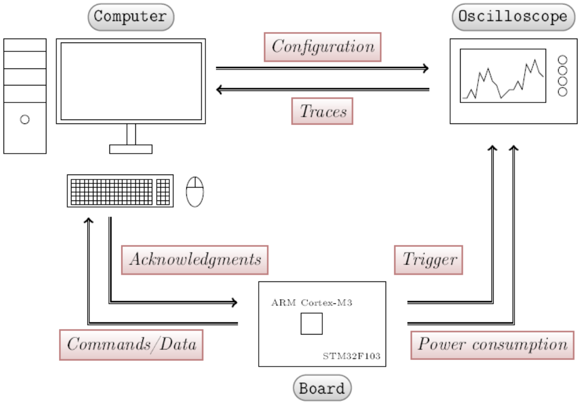
$$\tilde{C}_{p_i} = \underbrace{(s \;\&\; k)}_{\substack{\text{true only if } s=11\ldots1 \\ \text{else false}}} \quad | \quad \underbrace{((\sim s) \;\&\; \mu)}_{\substack{\text{true only if } s=00\ldots0 \\ \text{else false}}}$$

Giving:



Outline

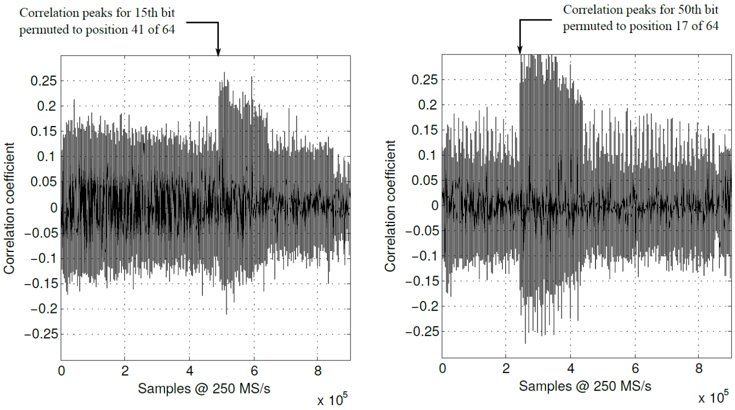
Attack bench [PRDCF16]



Traces analysis [PRDCF16]

- Apply a Hamming weight of individual bits leakage model: $H_i \in \{0, 1\}$,
- Use correlation coefficient to test our hypotheses compared with measurements,
- Good hypothesis if the coefficient is (almost) 1 or -1,
- Average of 500 traces per ciphertext hypothesis to avoid noise,
- Chosen ciphertexts as every vectors of weight 1.

Traces examples [PRDCF16]



Countermeasure [PRDCF16]

Algorithm

Inputs: Private permutation matrix $\mathcal{P}^{-1} \in \mathscr{M}_{n,n}(\mathbb{F}_2)$ represented by a lookup table $t^{\mathcal{P}^{-1}}$, ciphertext $\tilde{C} \in \mathbb{F}_2^n$ and private generator matrix $\mathcal{G}$ of $\Gamma(\mathscr{L}, G)$.

Randomly choose $B \in \Gamma(\mathscr{L}, G)$
$B_p = B \cdot \mathcal{P}$
$\tilde{C}' = \tilde{C} \oplus B_p$
For $i = 0$ **to** $n-1$
  $j = t_i^{\mathcal{P}^{-1}}$
  $\tilde{C}_{p_i}{}' = 0$
  For $h = 0$ **to** $n-1$
    $k = \tilde{C}_{p_i}{}'$
    $\mu = \tilde{C}_h{}'$
    $s = j \oplus h$
    $s \mid= s \gg 1$
    $s \mid= s \gg 2$
    $s \mid= s \gg 4$
    $s \mid= s \gg 8$
    $s \mid= s \gg 16$
    $s \mathrel{\&}= 1$
    $s = \sim (s-1)$
    $\tilde{C}_{p_i}{}' = (s \,\&\, k) \mid ((\sim s) \,\&\, \mu)$
  Endfor
Endfor
Return $\tilde{C}_p{}'$

Output: Permuted ciphertext $\tilde{C}_p' \in \mathbb{F}_2^n$ masked by a codeword.

Countermeasure [PRDCF16]

Main idea

From masked ciphertext to masked permuted ciphertext:

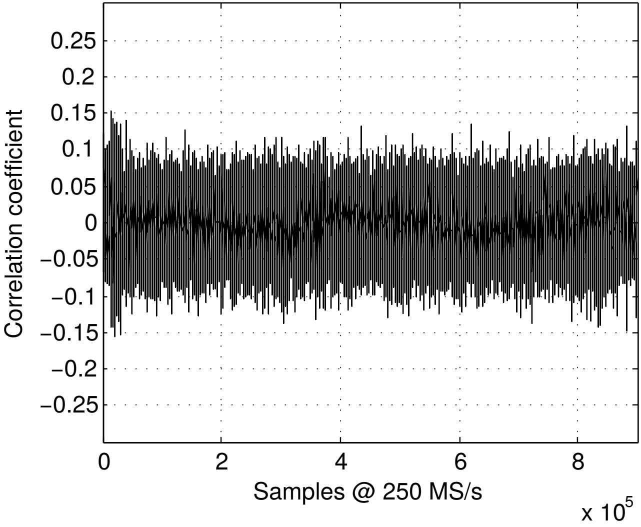
$$\begin{aligned}\tilde{C}'_p &= \tilde{C}' \cdot \mathcal{P}^{-1} \\ &= (\tilde{C} \oplus B_p) \cdot \mathcal{P}^{-1} \\ &= \tilde{C} \cdot \mathcal{P}^{-1} \oplus (B \cdot \mathcal{P}) \cdot \mathcal{P}^{-1} \\ &= \tilde{C}_p \oplus B.\end{aligned}$$

From masked permuted ciphertext to the same syndrome than non-masked ciphertext:

$$\begin{aligned}S &= \tilde{C}'_p \cdot \mathcal{H}^T \\ &= (\tilde{C}_p \oplus B) \cdot \mathcal{H}^T \\ &= \tilde{C}_p \cdot \mathcal{H}^T \oplus \underbrace{B \cdot \mathcal{H}^T}_{=0} \\ &= \tilde{C}_p \cdot \mathcal{H}^T.\end{aligned}$$

Countermeasure [PRDCF16]

Trace example



Outline

Context

Ciphertext permutation

DPA attack

Conclusion

Conclusion

- DPA against a 'secure' permutation algorithm (countermeasure for cache-memory attack),
- Simple masking countermeasure (with n more bits and not a huge amount of additional computations),
- DPA not depending on the code structure so possible for others linear codes than Goppa codes.

DPA on the 'Secure' Permutation in the McEliece PKC

Tania RICHMOND

Thank you for your attention!

UNIVERSITÉ DE TOULON

References

DH76 : *New directions in cryptography*, W. Diffie and M. Hellman, IEEE Transactions on Information Theory 1976.

McE78 : *A public-key cryptosystem based on algebraic coding theory*, R. McEliece, DSN progress report 1978.

STMOS08 : *Side Channels in the McEliece PKC*, F. Strenzke, E. Tews, H. G. Molter, R. Overbeck and A. Shoufan, PQCrypto 2008.

PRDCF16 : *Differential power analysis attack on the secure bit permutation in the McEliece cryptosystem*, M. Petrvalský, T. Richmond, M. Drutarovský, P.-L. Cayrel and V. Fischer, Radioelektronika 2016.

Pearson's correlation coefficient

We used for correlation analyses:

$$r_{H,X}(\eta) = \frac{\sum_{i=1}^{N}[(X_i(\eta) - \bar{X}(\eta))(H_i - \bar{H})]}{\sqrt{\sum_{i=1}^{N}[X_i(\eta) - \bar{X}(\eta)]^2 \sum_{i=1}^{N}(H_i - \bar{H})^2}}$$

where $r_{H,X}(\eta)$ is the Pearson's correlation coefficient for η-th sample (measured during execution of the cryptographic algorithm), N is a number of measured traces, $X_i(\eta)$ is a value of η-th sample measured during i-th measurement (i-th trace), $\bar{X}(\eta)$ is a mean value of corresponding η-th samples (from all traces), H_i is a hypothesis of power consumption for one bit of input data corresponding with i-th measurement (i-th trace) and $\bar{H}$ is a mean value of all hypotheses H_i.