Side-Channel Analysis of Multiplications in $GF(2^{128})$

Application to AES-GCM

Sonia Belaïd[1] Pierre-Alain Fouque[2] Benoît Gérard[3]

[1]École normale supérieure and Thales Communications & Security,

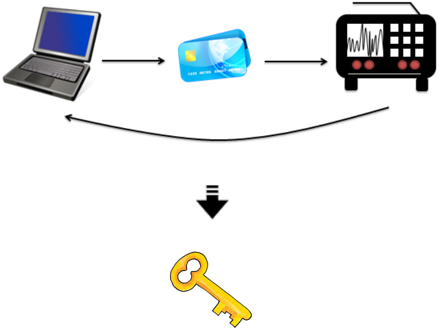
[2]Université de Rennes 1 and Institut Universitaire de France

[3]DGA–MI and IRISA

11-26-2014

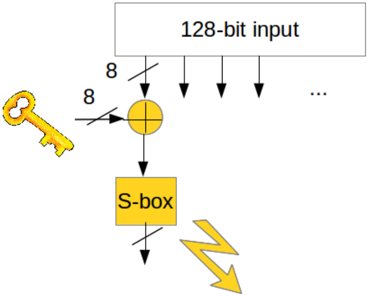
Side-Channel Attacks

- physical leakage
  - timing
  - power consumption
  - temperature
  - ...
- statistical treatment
- key recovery



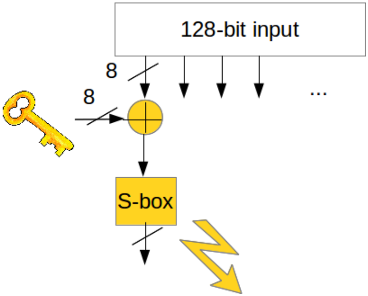
Key-Dependent Leakage

AES Block Cipher

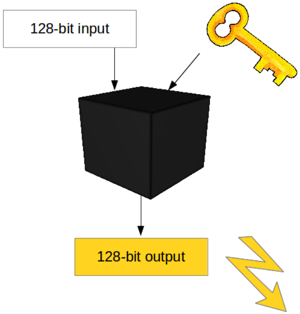


Key-Dependent Leakage

AES Block Cipher



GCM : Finite Field Multiplication

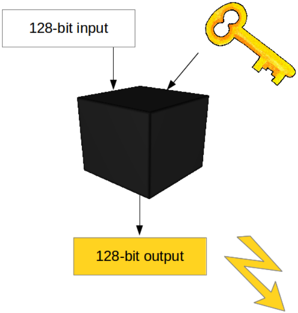


Contributions

Side-Channel Analysis of Multiplications in GF(2^{128}) : Application to AES-GCM

Sonia Belaïd, Pierre-Alain Fouque, Benoît Gérard

Asiacrypt 2014



Outline

Outline

1. Target Primitive : AES-GCM
2. Attack
   - Main Idea
   - Known Inputs
   - Chosen Inputs
3. Extensions
   - Another Application: Re-keying
   - Specific Implementations
4. Conclusion

AES-GCM

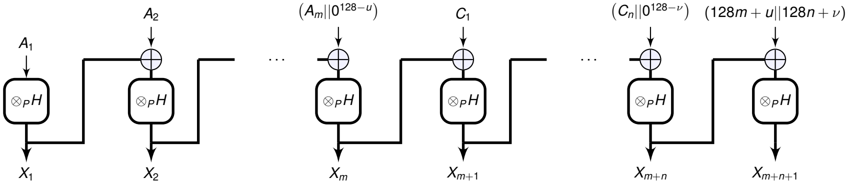


Figure: AES-GCM authentication

hashed key H: $H = \text{AES}_K(0^{128})$ with K the encryption key
authenticated data A_i: 128-bit blocks of data to authenticate
ciphertexts C_i: 128-bit encrypted blocks

Galois Field Multiplication $\otimes_P$

$$\mathrm{GF}(2^{128}) = \mathrm{GF}(2)/P(Y), \quad P(Y) = Y^{128} + Y^7 + Y^2 + Y + 1$$

$$M_P \otimes_P H =$$

$$\begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_{127} \end{pmatrix}$$

$$\begin{pmatrix} m_0 & m_{127} & \cdots & m_1 \oplus m_{127} \oplus m_{126} \\ m_1 & m_0 \oplus m_{127} & \cdots & m_2 \oplus m_{123} \oplus m_1 \oplus m_{127} \oplus m_{122} \\ \vdots & \vdots & \ddots & \vdots \\ m_{127} & m_{126} & \cdots & m_0 \oplus m_{127} \oplus m_{126} \oplus m_{121} \end{pmatrix} \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix}$$

Galois Field Multiplication $\otimes_P$

$$\mathrm{GF}(2^{128}) = \mathrm{GF}(2)/P(Y), \quad P(Y) = Y^{128} + Y^7 + Y^2 + Y + 1$$

$$M_P \otimes_P H =$$

$$\begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_{127} \end{pmatrix}$$

$$\begin{pmatrix} m_0 & m_{127} & \cdots & m_1 \oplus m_{127} \oplus m_{126} \\ m_1 & m_0 \oplus m_{127} & \cdots & m_2 \oplus m_{123} \oplus m_1 \oplus m_{127} \oplus m_{122} \\ \vdots & \vdots & \ddots & \vdots \\ m_{127} & m_{126} & \cdots & m_0 \oplus m_{127} \oplus m_{126} \oplus m_{121} \end{pmatrix} \begin{pmatrix} x_0 \\ \\ \\ \\ \end{pmatrix}$$

Galois Field Multiplication $\otimes_P$

$$\mathrm{GF}(2^{128}) = \mathrm{GF}(2)/P(Y), \quad P(Y) = Y^{128} + Y^7 + Y^2 + Y + 1$$

$$M_P \otimes_P H =$$

$$\begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_{127} \end{pmatrix}$$

$$\begin{pmatrix} m_0 & m_{127} & \cdots & m_1 \oplus m_{127} \oplus m_{126} \\ m_1 & m_0 \oplus m_{127} & \cdots & m_2 \oplus m_{123} \oplus m_1 \oplus m_{127} \oplus m_{122} \\ \vdots & \vdots & \ddots & \vdots \\ m_{127} & m_{126} & \cdots & m_0 \oplus m_{127} \oplus m_{126} \oplus m_{121} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \\ \\ \end{pmatrix}$$

Galois Field Multiplication $\otimes_P$

$$\mathrm{GF}(2^{128}) = \mathrm{GF}(2)/P(Y), \quad P(Y) = Y^{128} + Y^7 + Y^2 + Y + 1$$

$$M_P \otimes_P H =$$

$$\begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_{127} \end{pmatrix}$$

$$\begin{pmatrix} m_0 & m_{127} & \cdots & m_1 \oplus m_{127} \oplus m_{126} \\ m_1 & m_0 \oplus m_{127} & \cdots & m_2 \oplus m_{123} \oplus m_1 \oplus m_{127} \oplus m_{122} \\ \vdots & \vdots & \ddots & \vdots \\ m_{127} & m_{126} & \cdots & m_0 \oplus m_{127} \oplus m_{126} \oplus m_{121} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{127} \end{pmatrix}$$

Outline

Leakage Models

Hamming Weight

$$L_i^{(\mathrm{HW})} = \mathrm{HW}(V_i) + \varepsilon_\sigma, \quad \varepsilon_\sigma \sim \mathcal{N}(0, \sigma)$$

Hamming Distance

$$L_i^{(\mathrm{HD})} = \mathrm{HD}(V_i, V_{i-1}) + \varepsilon_\sigma = \mathrm{HW}(V_i \oplus V_{i-1}) + \varepsilon_\sigma$$

Attacker Capabilities

Known/Chosen Inputs:

- ciphertexts
- authenticated data

Limited/Unlimited Queries:

- error-counter for the tag verifications

Enabled/Disabled Averaging:

- specific formats
- same computation λ times: $\sigma \mapsto \sigma/\sqrt{\lambda}$

Outline

1. Target Primitive : AES-GCM
2. Attack
   - Main Idea
   - Known Inputs
   - Chosen Inputs
3. Extensions
   - Another Application: Re-keying
   - Specific Implementations
4. Conclusion

Main Idea of The Attack

Current Issue: each bit of the multiplication of two 128-bit blocks depends on all the input bits

- ✘ no divide-and-conquer strategy

Main observation: the LSB of the Hamming Weight (same for HD) of a variable is a linear function of its bits:

$$\mathrm{lsb}_0\left(\mathrm{HW}(V)\right) = \bigoplus_{0 \leqslant i \leqslant 127} v_i$$

Main Idea of The Attack

Current Issue: each bit of the multiplication of two 128-bit blocks depends on all the input bits

- ✘ no divide-and-conquer strategy

Main observation: the LSB of the Hamming Weight (same for HD) of a variable is a linear function of its bits:

$$\text{lsb}_0\,(\text{HW}(V)) = \bigoplus_{0 \leqslant i \leqslant 127} v_i$$

- $$\text{lsb}_0\,(\text{HW}(M \otimes_P H)) = \bigoplus_{0 \leqslant i \leqslant 127} \left(\bigoplus_{0 \leqslant j \leqslant 127} (M \otimes_P \alpha^i)_j \right) h_i$$

New Issue

New Issue: leakage comes with noise

$$\begin{aligned}\widetilde{b_0} &\stackrel{\text{def}}{=} \text{lsb}_0\left(\lceil \text{HW}(M \otimes_P H) + \varepsilon_\sigma \rfloor\right) \\ &= \text{lsb}_0\left(\text{HW}(M \otimes_P H)\right) \oplus b_{\mathcal{N}}\end{aligned}$$

$b_{\mathcal{N}}$ follows a Bernoulli distribution with a parameter p such that the probability of no error is

$$1 - p = \sum_{i=-\infty}^{\infty} \int_{2i-0.5}^{2i+0.5} e^{-\frac{t^2}{2\sigma^2}} / (\sigma\sqrt{2\pi}) dt$$

σ	0.5	1	2	3	4	5
p	0.31	$0.5 - 4.6.10^{-3}$	$0.5 - 1.7.10^{-9}$	$0.5 - \varepsilon$	$0.5 - \varepsilon$	$0.5 - \varepsilon$

Application on the other bits ?

$$b_i = \bigoplus_{0 \leqslant j_1 < \cdots < j_{2^i} \leqslant 127} \left(\prod_{1 \leqslant \ell \leqslant 2^i} \bigoplus_{0 \leqslant k \leqslant 127} (M \otimes_P \alpha^k)_{j_\ell} \; h_k \right), \;\; \forall\, 0 \leqslant i \leqslant 7$$

σ	Bernoulli parameter p							
	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7
0.5	$3.1\ 10^{-1}$	$1.6\ 10^{-1}$	$8.0\ 10^{-2}$	$4.0\ 10^{-2}$	$2.3\ 10^{-2}$	$2.2\ 10^{-2}$	$2.2\ 10^{-2}$	ε
1	$\frac{1}{2} - 4.6\ 10^{-3}$	$3.7\ 10^{-1}$	$1.9\ 10^{-1}$	$9.5\ 10^{-2}$	$5.5\ 10^{-2}$	$5.3\ 10^{-2}$	$5.3\ 10^{-2}$	ε
2	$\frac{1}{2} - 1.5\ 10^{-4}$	$\frac{1}{2} - 3.2\ 10^{-3}$	$3.8\ 10^{-1}$	$2.0\ 10^{-1}$	$1.1\ 10^{-1}$	$1.1\ 10^{-1}$	$1.1\ 10^{-1}$	ε
3	$\frac{1}{2} - \varepsilon$	$\frac{1}{2} - 6.8\ 10^{-8}$	$4.7\ 10^{-1}$	$3.0\ 10^{-1}$	$1.6\ 10^{-1}$	$1.5\ 10^{-1}$	$1.5\ 10^{-1}$	ε
4	$\frac{1}{2} - \varepsilon$	$\frac{1}{2} - 1.2\ 10^{-9}$	$\frac{1}{2} - 3.0\ 10^{-3}$	$3.8\ 10^{-1}$	$2.1\ 10^{-1}$	$1.9\ 10^{-1}$	$1.9\ 10^{-1}$	ε
5	$\frac{1}{2} - \varepsilon$	$\frac{1}{2} - \varepsilon$	$\frac{1}{2} - 1.9\ 10^{-4}$	$4.4\ 10^{-1}$	$2.6\ 10^{-1}$	$2.3\ 10^{-1}$	$2.3\ 10^{-1}$	ε

Outline

Naive Attack

$$\mathcal{S} = \begin{cases} \bigoplus_{0 \leqslant i \leqslant 127} \left(\bigoplus_{0 \leqslant j \leqslant 127} (M^{(0)} \otimes_P \alpha^i)_j \right) h_i = \widetilde{b_0}^{(0)} \\ \bigoplus_{0 \leqslant i \leqslant 127} \left(\bigoplus_{0 \leqslant j \leqslant 127} (M^{(1)} \otimes_P \alpha^i)_j \right) h_i = \widetilde{b_0}^{(1)} \\ \qquad \ldots \\ \bigoplus_{0 \leqslant i \leqslant 127} \left(\bigoplus_{0 \leqslant j \leqslant 127} (M^{(t-1)} \otimes_P \alpha^i)_j \right) h_i = \widetilde{b_0}^{(t-1)} \end{cases}$$

To solve the system, two conditions must be fulfilled:

i) $\mathcal{S}$ contains at least as many linearly independent equations as the number of unknown variables (128),

ii) there is no error in the bits $\widetilde{b_0}^{(\ell)}$ (i.e., $\widetilde{b_0}^{(\ell)} = b_0^{(\ell)}$).

Naive Attack

i) $\mathcal{S}$ contains at least 128 linearly independent equations,

- probability from 128 messages ≈ 0.3
- probability from 129 messages ≈ 0.9

Naive Attack

i) $\mathcal{S}$ contains at least 128 linearly independent equations,
   - probability from 128 messages ≈ 0.3
   - probability from 129 messages ≈ 0.9

ii) there is no error in the bits $\widetilde{b_0}^{(\ell)}$ (i.e., $\widetilde{b_0}^{(\ell)} = b_0^{(\ell)}$)
   - complexity to remove e errors:

$$C_{128}^{(e)} = \sum_{i=0}^{e} \binom{128}{i}$$

   - if $e = 6$, $C_{128}^{(e)} \approx 2^{32}$

Improved Attack

- Reducing the Noise Impact
- Saving Traces
- Solving the System with more Errors and Advanced Algorithms

An Optimal Decision Rule

Idea: use the LLR (Log Likelihood Ratio) to approximate better the bit value b_0

$$\widehat{b_0} \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } \mathrm{LLR}(\ell) \geqslant 0, \\ 1 & \text{otherwise.} \end{cases}$$

with

$$\mathrm{LLR}(\ell) = \log(\ \mathbb{P}[b_0 = 0 \mid \ell]\ ) - \log(\ \mathbb{P}[b_0 = 1 \mid \ell]\ )$$

instead of

$$\widetilde{b_0} \stackrel{\text{def}}{=} \mathrm{lsb}_0\left(\lceil \mathrm{HW}(M \otimes_P H) + \varepsilon_\sigma \rfloor\right)$$

Selecting Traces

When more than 128 traces are available,

Idea: choose 128 linearly independent samples from the highest LLR values

Selecting Traces

When more than 128 traces are available,

- *Idea:* choose 128 linearly independent samples from the highest LLR values
- *Issue:* maximizing this LLR sum is a combinatorial optimization problem quite hard to solve

Selecting Traces

When more than 128 traces are available,

- *Idea:* choose 128 linearly independent samples from the highest LLR values
- *Issue:* maximizing this LLR sum is a combinatorial optimization problem quite hard to solve
- *Solution:* first come/first selected algorithm: iteratively pick the highest LLR value which increases the system rank

Reducing the Noise: Selecting Traces

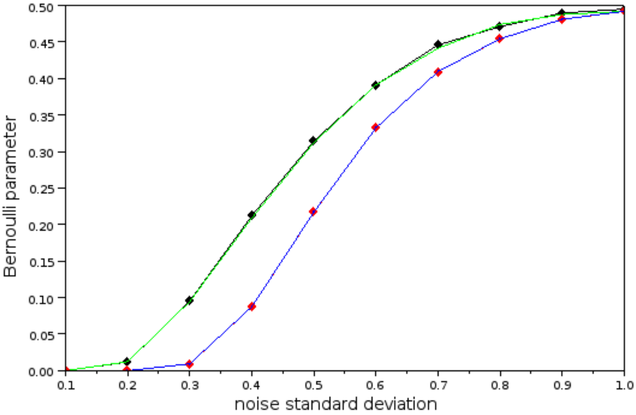


Figure: Bernoulli parameter with rounding (black), LLR (green), traces selection (blue) and best LLR traces (red)

Saving Traces

In AES-GCM,

$$\begin{aligned} X_2 &= (M_1 \otimes_P H \oplus M_2) \otimes_P H \\ &= M_1 \otimes_P H^2 \oplus M_2 \otimes_P H \end{aligned}$$

Since squaring is linear over GF(2), there exists S such that

$$X_2 = (M_1 \cdot S \oplus M_2) \otimes_P H$$

- two multiplications with a single trace

Solving the System with more Errors and Advanced Algorithms

Noisy codeword: LSBs extracted from leaking multiplications that encode the authentication key H

Issue: decoding the noisy codeword

- Learning Parities with Noise (LPN) Algorithms
- Linear Decoding

Method \ σ	0.1 C_s/C_t	0.2 C_s/C_t	0.3 C_s/C_t	0.4 C_s/C_t	0.5 C_s/C_t
LLR + naive	$2^8/2^{21}$	$2^8/2^{21}$	$2^8/2^{22}$	$2^8/2^{65}$	$2^8/2^{107}$
LPN (LF Algo)	$2^{11}/2^{14}$	$2^{20}/2^{22}$	$2^{26}/2^{28}$	$2^{32}/2^{34}$	$2^{48}/2^{50}$
Linear decoding	$2^6/2^6$	$2^6/2^7$	$2^7/2^{11}$	$2^8/2^{25}$	$2^9/2^{62}$

Outline

Improvements

If the attacker can choose the messages, some improvements are:

- averaging the traces
- structuring the messages to make the system easier to solve
- choosing messages to be able to exploit more than two multiplications

Averaging Traces

Repeating the same computation λ times: $\sigma \mapsto \sigma/\sqrt{\lambda}$

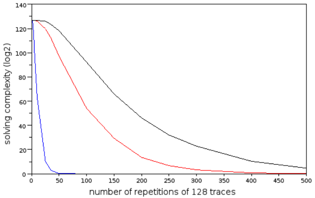


Figure: Solving complexities with repetitions for $\sigma = 1$ (blue), $\sigma = 3$ (red) and $\sigma = 4$ (black)

Experimental Results: tests on the Virtex-5 FPGA of a SASEBO board with an EM probe for the acquisition

- confirm the simulations

Structuring the Messages

Which code should we use?

- *List decoding:* concatenation of smaller linear codes to recover the key chunks

► concatenated code of smaller random linear codes which can be efficiently decoded using a Fast Walsh Transform

$$\begin{pmatrix} \boxed{\mathcal{S}_0} & & \\ & \boxed{\mathcal{S}_1} & \\ & & \ddots \end{pmatrix} \cdot \begin{pmatrix} H \end{pmatrix} = \begin{pmatrix} \widehat{b_0} \\ \vdots \\ \widehat{b_t} \end{pmatrix}$$

Structuring the Messages

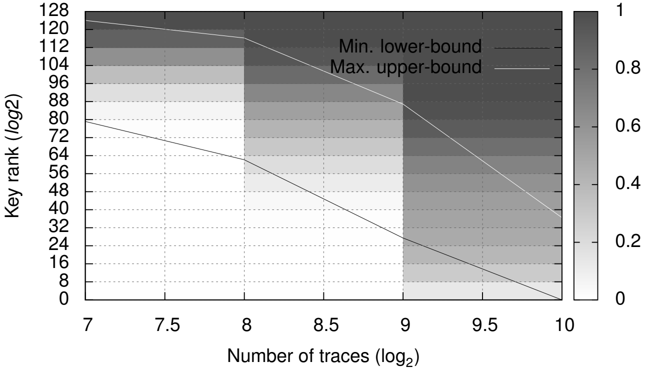


Figure: Security graph for $\sigma = 0.5$

Saving Traces

Saving Traces: exploit the linearity of the squaring operation (as suggested by Ferguson)

$$
\begin{aligned}
X_1 &= M_1 \otimes_P H, \\
X_2 &= M_1 \otimes_P H^2 \oplus M_2 \otimes_P H, \\
X_3 &= M_1 \otimes_P H^3 \oplus M_2 \otimes_P H^2 \oplus M_3 \otimes_P H, \\
X_4 &= M_1 \otimes_P H^4 \oplus M_2 \otimes_P H^3 \oplus M_3 \otimes_P H^2 \oplus M_4 \otimes_P H.
\end{aligned}
$$

Saving Traces

Saving Traces: exploit the linearity of the squaring operation (as suggested by Ferguson)

$$
\begin{aligned}
X_1 &= M_1 \otimes_P H, \\
X_2 &= M_1 \otimes_P H^2 \oplus M_2 \otimes_P H, \\
X_3 &= M_1 \otimes_P H^3 \oplus M_2 \otimes_P H^2 \oplus M_3 \otimes_P H, \\
X_4 &= M_1 \otimes_P H^4 \oplus M_2 \otimes_P H^3 \oplus M_3 \otimes_P H^2 \oplus M_4 \otimes_P H.
\end{aligned}
$$

$$M_2 = 0$$

Saving Traces

Saving Traces: exploit the linearity of the squaring operation (as suggested by Ferguson)

$$
\begin{aligned}
&\blacktriangleright \quad X_1 &=& \quad M_1 \otimes_P H, \\
&\blacktriangleright \quad X_2 &=& \quad M_1 \otimes_P H^2, \\
&\quad\;\; X_3 &=& \quad M_1 \otimes_P H^3 \oplus M_3 \otimes_P H, \\
&\blacktriangleright \quad X_4 &=& \quad M_1 \otimes_P H^4 \oplus M_3 \otimes_P H^2 \oplus M_4 \otimes_P H.
\end{aligned}
$$

$$M_2 = 0$$

Outline

Outline

Re-keying from Medwed et al.[1]

$$k^\star = r \cdot k \quad \in \quad \mathrm{GF}(2^8)[Y]/P(Y) = Y^{16} + 1$$

- in the matrix/vector product $K^\star = R_P \otimes_P K$:

$$R_p = \begin{pmatrix} r_0 & r_{15} & \cdots & r_1 \\ r_1 & r_0 & \cdots & r_2 \\ \vdots & \vdots & \ddots & \vdots \\ r_{15} & r_{14} & \cdots & r_0 \end{pmatrix}$$

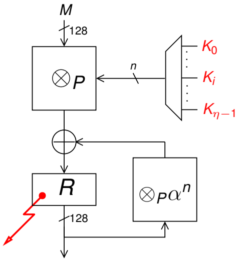
- equation of the LSB:

$$\mathrm{lsb}_0 \left(\mathrm{HW} \left[\left(\bigoplus_{0 \leqslant i \leqslant m-1} r_i \right) \cdot \left(\bigoplus_{0 \leqslant j \leqslant m-1} k_j \right) \right] \right) = b_0$$

[1] M. Medwed, C. Petit, F. Regazzoni, M. Renauld, F.-X. Standaert, Fresh Re-Keying II: Securing Multiple Parties against Side-Channel and Fault Attacks, CARDIS 2011

Outline

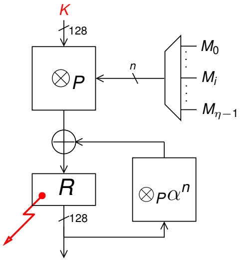
Specific Implementations



if the key is split

- divide-and-conquer strategy

Specific Implementations



if the key is split

- divide-and-conquer strategy

---

if the message is split

- sparse messages
- easier than the generic (known inputs) scenario

Outline

Conclusion

- Summary
  - attack the AES-GCM authentication without looking inside the multiplication
  - exploitation of the LSB
  - different improvements
- Further Work
  - application of similar attacks to other primitives
  - exploitation of more leakage bits with different techniques

Thank you

Thank you for your attention.