Cryptanalysis of the counter mode of operation

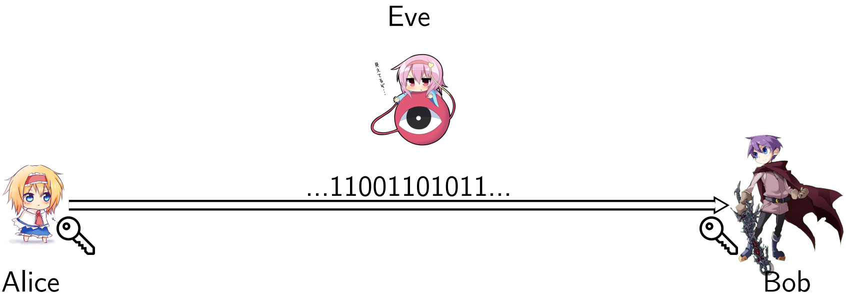
Ferdinand Sibleyras
joint work with Gaëtan Leurent

Inria, équipe SECRET

April 10, 2018

Introduction

- **Cryptography:** Alice encrypts then sends messages to Bob.
- **Symmetric:** Alice and Bob share the same key.
- **Public channel:** Eve (attacker) can see and/or manipulate what is being sent.



Introduction

Block Cipher

$$E_k : \{0,1\}^n \rightarrow \{0,1\}^n$$

A family of **permutations** indexed by a key (AES, 3DES, ...) where n is the bit size of the permutation or block's size.

Introduction

Block Cipher

$$E_k : \{0,1\}^n \to \{0,1\}^n$$

A family of **permutations** indexed by a key (AES, 3DES, ...) where n is the bit size of the permutation or block's size.

Mode of operation

Describes how to use a **block cipher** along with a plaintext message of **arbitrary length** to achieve some concrete cryptographic goals.

Introduction

Modes are classified according to their goals:

- There are encryption modes (CBC, CTR, ...).
  They aim at hiding the plaintext.
  $\rightarrow$ Plaintext recovery attacks.

Introduction

Modes are classified according to their goals:

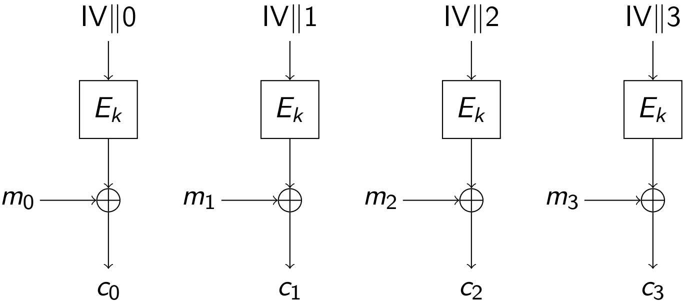
- There are encryption modes (CBC, CTR, ...).
  They aim at hiding the plaintext.
  $\rightarrow$ Plaintext recovery attacks.
- There are authentication modes (GMAC, ...).
  They aim at authenticating the plaintext.
  $\rightarrow$ Forgery attacks.

Introduction

Modes are classified according to their goals:

- There are encryption modes (CBC, CTR, ...).
  They aim at hiding the plaintext.
  $\rightarrow$ Plaintext recovery attacks.
- There are authentication modes (GMAC, ...).
  They aim at authenticating the plaintext.
  $\rightarrow$ Forgery attacks.
- There are authenticated encryption modes (GCM, ...).
  They aim at both authenticating and hiding the plaintext.

The counter mode (CTR)



m_i : The plaintext.

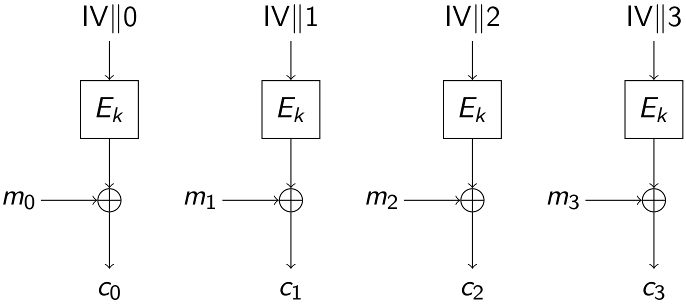
c_i : The ciphertext.

E_k : The block cipher.

IV : The Initialisation Value.

$$c_i = E_k(\mathrm{IV} \| i) \oplus m_i$$

The counter mode (CTR)

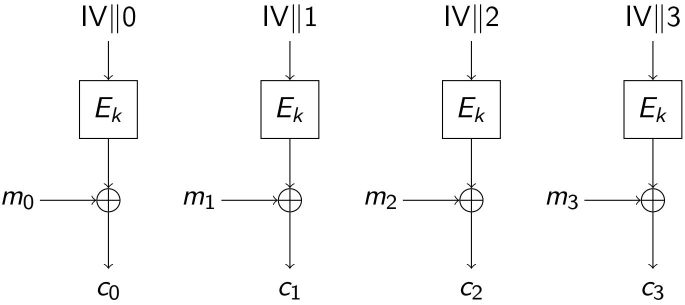


m_i : The plaintext. E_k : The block cipher.
c_i : The ciphertext. IV : The Initialisation Value.

$$c_i = E_k(\mathrm{IV} \| i) \oplus m_i$$

Akin to a stream cipher: keystream XORed with the plaintext.

The counter mode (CTR)



m_i : The plaintext. E_k : The block cipher.
c_i : The ciphertext. IV : The Initialisation Value.

$$c_i = E_k(\mathrm{IV} \| i) \oplus m_i$$

Akin to a stream cipher: keystream XORed with the plaintext.
Inputs $\mathrm{IV}\|i$ to the block cipher never repeat.

The counter mode (CTR)

Let $K_i = E_k(\mathrm{IV} \| i)$ the ith block of keystream.

- If E_k is a good Pseudo-Random Function (PRF) then all K_i are random and this is a one-time-pad.
- A block cipher is a Pseudo-Random Permutation (PRP) therefore K_i are all distinct: $K_i \neq K_j \; \forall i \neq j$.

The counter mode (CTR)

Let $K_i = E_k(\mathrm{IV}\|i)$ the ith block of keystream.

- If E_k is a good Pseudo-Random Function (PRF) then all K_i are random and this is a one-time-pad.
- A block cipher is a Pseudo-Random Permutation (PRP) therefore K_i are all distinct: $K_i \neq K_j \ \forall i \neq j$.

Security proof (σ the number of blocks)

$$\mathbf{Adv}^{\mathrm{CPA}}_{\mathrm{CTR}\text{-}E_k}(\sigma) \leq \mathbf{Adv}^{\mathrm{PRF}}_{E_k}(\sigma) \leq \mathbf{Adv}^{\mathrm{PRP}}_{E_k}(\sigma) + \sigma^2/2^{n+1}$$

Distinguishing attack

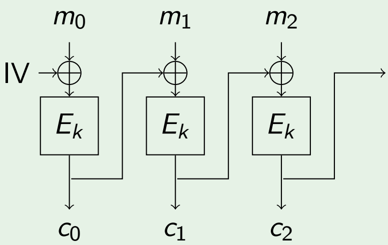
After $\sigma \simeq 2^{n/2}$ encrypted blocks we expect a collision on the K_i with high probability in the case of a random ciphertext.
That is the birthday bound coming from the birthday paradox.

CBC and CTR

Both modes are:

- widely deployed
- proven secure up to birthday bound ($2^{n/2}$)
- allowing attacks when nearing the bound

CBC mode

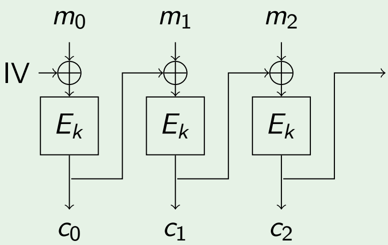


CBC and CTR

Both modes are:

- widely deployed
- proven secure up to birthday bound ($2^{n/2}$)
- allowing attacks when nearing the bound

CBC mode



Folklore assumptions [Ferguson, Schneier, Kohno]

CTR leaks very little data. [...] It would be reasonable to limit the cipher mode to 2^{60} blocks, which allows you to encrypt 2^{64} bytes but restricts the leakage to a small fraction of a bit.
When using CBC mode you should be a bit more restrictive. [...] We suggest limiting CBC encryption to 2^{32} blocks or so.

The counter mode (CTR)

From a distinguishing attack to a plaintext recovery attack ?

- If we know m_i, we recover $K_i = c_i \oplus m_i$.

The counter mode (CTR)

From a distinguishing attack to a plaintext recovery attack ?

- If we know m_i, we recover $K_i = c_i \oplus m_i$.
- We can observe repeated encryptions of a secret S that is $c_j = K_j \oplus S$ for many different j.

The counter mode (CTR)

From a distinguishing attack to a plaintext recovery attack ?

- If we know m_i, we recover $K_i = c_i \oplus m_i$.
- We can observe repeated encryptions of a secret S that is $c_j = K_j \oplus S$ for many different j.
- The distinguishing attack uses $K_i \oplus K_j \neq 0$ which implies $K_i \oplus c_j \neq S$ $\forall i \neq j$.

The counter mode (CTR)

From a distinguishing attack to a plaintext recovery attack ?

- If we know m_i, we recover $K_i = c_i \oplus m_i$.
- We can observe repeated encryptions of a secret S that is $c_j = K_j \oplus S$ for many different j.
- The distinguishing attack uses $K_i \oplus K_j \neq 0$ which implies $K_i \oplus c_j \neq S$ $\forall i \neq j$.

Main Idea

Collect many keystream blocks K_i and encryptions of secret block $c_j = K_j \oplus S$; then look for a value s such that $K_i \oplus c_j \neq s$ $\forall i \neq j$.

Missing difference problem

The missing difference problem

- Given $\mathcal{A}$ and $\mathcal{B}$, and a hint $\mathcal{S}$ three sets of n-bit words
- Find $S \in \mathcal{S}$ such that:

$$\forall (a, b) \in \mathcal{A} \times \mathcal{B}, \; S \neq a \oplus b \; .$$

Missing difference problem

Main Idea

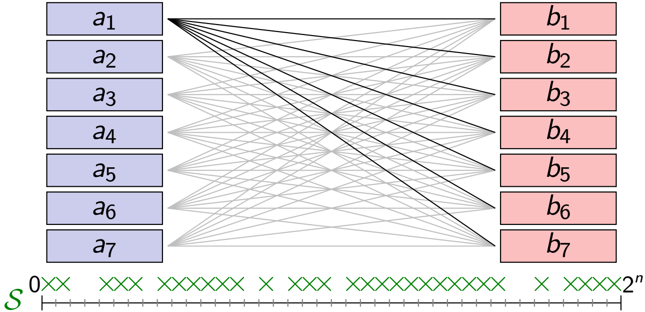
Collect many keystream blocks $K_i \in \mathcal{A}$ and encryptions of secret block $c_j = K_j \oplus S \in \mathcal{B}$; then look for a value $s \in \mathcal{S}$ such that $\forall (a, b) \in \mathcal{A} \times \mathcal{B},\ s \neq a \oplus b$.

The missing difference problem

- Given $\mathcal{A}$ and $\mathcal{B}$, and a hint $\mathcal{S}$ three sets of n-bit words
- Find $S \in \mathcal{S}$ such that:

$$\forall (a, b) \in \mathcal{A} \times \mathcal{B},\ S \neq a \oplus b\ .$$

Simple Sieving Algorithm [McGrew, FSE'13]

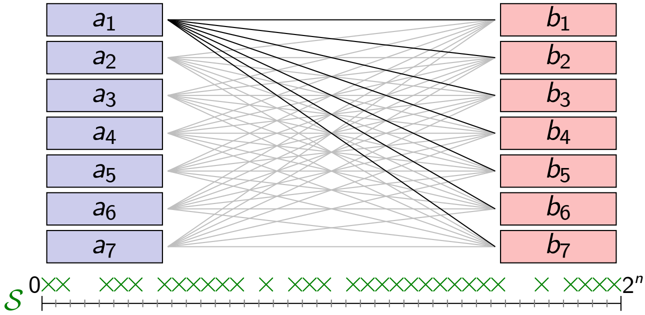


Compute all $a_i \oplus b_j$, remove results from a sieve $\mathcal{S}$.

Analysis: case $|\mathcal{S}| = 2^n$ via coupon collector problem

- To exclude 2^n candidates of S, we need $n \cdot 2^n$ values $a_i \oplus b_j$
  - Lists $\mathcal{A}$ and $\mathcal{B}$ of size $\sqrt{n} \cdot 2^{n/2}$. Complexity: $\tilde{\mathcal{O}}(2^n)$

Simple Sieving Algorithm [McGrew, FSE'13]

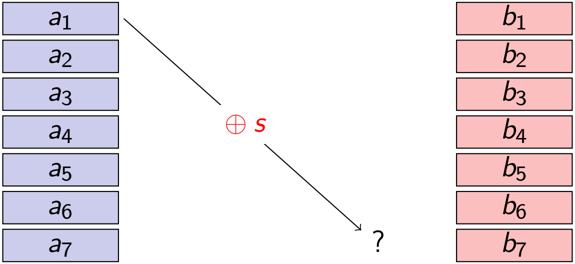


Compute all $a_i \oplus b_j$, remove results from a sieve $\mathcal{S}$.

Analysis: case $|\mathcal{S}| = 2$

- To exclude 1 candidate of S, we need 2^n values $a_i \oplus b_j$
  - Lists $\mathcal{A}$ and $\mathcal{B}$ of size $2^{n/2}$. Complexity: $\tilde{\mathcal{O}}(2^n)$

Searching Algorithm [McGrew, FSE'13]

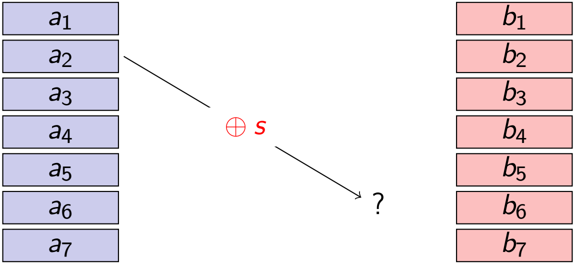


- Make a guess and verify.

Try Guess (s)

for a **in** $\mathcal{A}$ **do**
  if $(s \oplus a) \in \mathcal{B}$ **then**
    return 0
return 1

Searching Algorithm [McGrew, FSE'13]

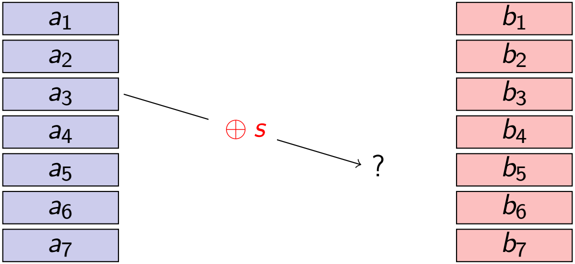


- Make a guess and verify.

Try Guess (s)

for a **in** $\mathcal{A}$ **do**
    if $(s \oplus a) \in \mathcal{B}$ **then**
        return 0
return 1

Searching Algorithm [McGrew, FSE'13]



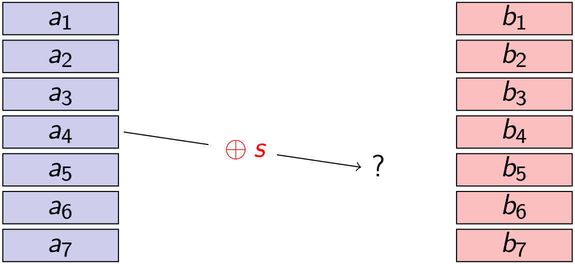
- Make a guess and verify.

Try Guess (s)

for a **in** $\mathcal{A}$ **do**
  if $(s \oplus a) \in \mathcal{B}$ **then**
    return 0
return 1

Searching Algorithm

[McGrew, FSE'13]



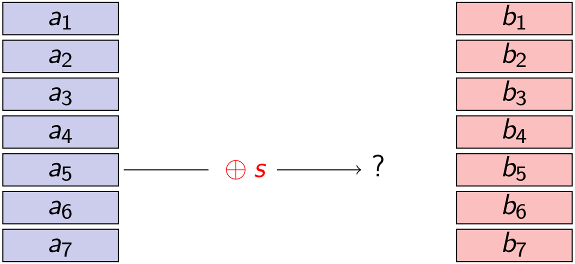
- Make a guess and verify.

Try Guess (s)

for a **in** $\mathcal{A}$ **do**
    if $(s \oplus a) \in \mathcal{B}$ **then**
        return 0
return 1

Searching Algorithm

[McGrew, FSE'13]

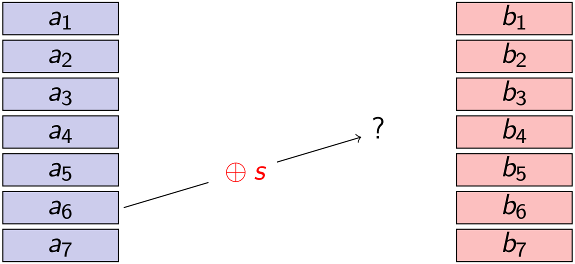


- Make a guess and verify.

Try Guess (s)

for a **in** $\mathcal{A}$ **do**
    if $(s \oplus a) \in \mathcal{B}$ **then**
        return 0
return 1

Searching Algorithm [McGrew, FSE'13]

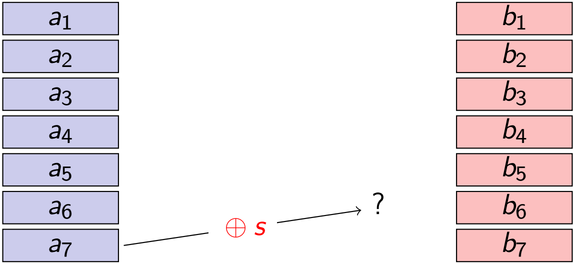


- Make a guess and verify.

Try Guess (s)

for a **in** $\mathcal{A}$ **do**
  if $(s \oplus a) \in \mathcal{B}$ **then**
    return 0
return 1

Searching Algorithm [McGrew, FSE'13]

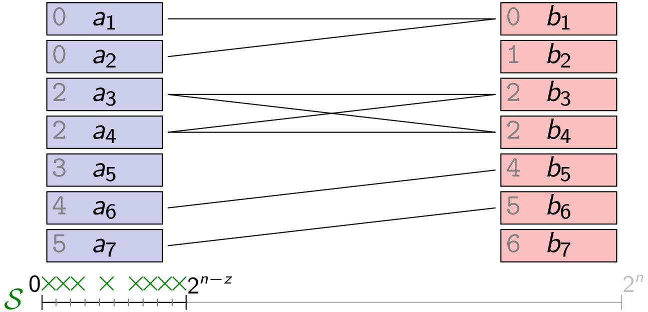


- Make a guess and verify.
- Complexity $\tilde{\mathcal{O}}(2^{n/2}\sqrt{|\mathcal{S}|})$ with unbalanced $\mathcal{A}$, $\mathcal{B}$.

Try Guess (s)

for a **in** $\mathcal{A}$ **do**
  if $(s \oplus a) \in \mathcal{B}$ **then**
    return 0
return 1

Known-prefix Sieving



- Assume S starts with z zero bits (more generally, linear subspace with $\dim\langle\mathcal{S}\rangle = n - z$)
- Sort lists, consider a_i's and b_j's with matching z-bit prefix
- Complexity: $\tilde{\mathcal{O}}(2^{n/2} + 2^{\dim\langle\mathcal{S}\rangle})$
  - Looking for collision + needed number of collisions
- Complexity: $\tilde{\mathcal{O}}(2^{n/2})$ when $\dim\langle\mathcal{S}\rangle \leq n/2$

Simulation

We challenge the false assumptions we made like independence of the $\{a \oplus b\}$. Approximations seem good enough.

Ran simulations with $n = 64$ bits and $z = n/2 = 32$ zeros.

- Each round we compare two lists of $2^{n/2}$ elements.
- Each round we expect $2^{n/2}$ partial collisions.
- Coupon collector predicts $n/2 \cdot \ln(2) \cdot 2^{n/2}$ partial collisions to recover S, that is 23 rounds on expectation.
- Simulation gives an idea of what is hidden in the $\mathcal{O}$ notations.

Consistent speed of leaking

In every runs, after 16 rounds the sieve was left between 419 and 560 candidates of S only.

Simulation

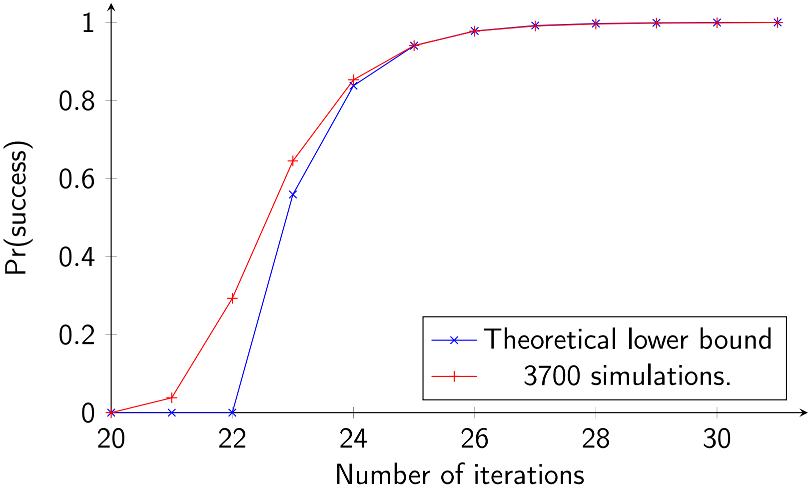
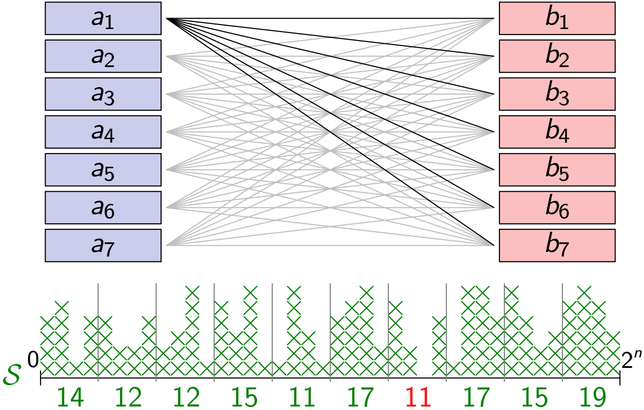


Figure: Probability of success of the known prefix sieving knowing 2^{32} encryptions of a 32-bit secret against the number of chunks of 2^{32} keystream blocks of size $n = 64$ bits used.

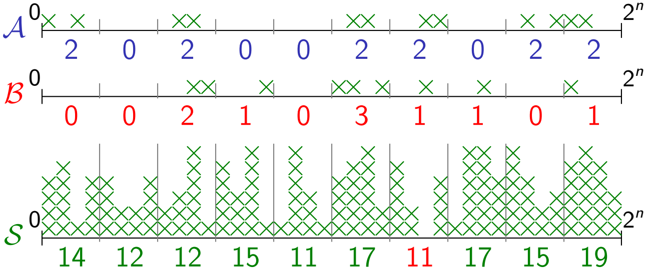
Fast Convolution Sieving



- Instead of computing full sieve, use buckets (*ie.* truncate)
- With enough data, missing difference has smallest bucket with high probability
  - *Eg.* $2^{2n/3}$ queries, sieving with $2^{2n/3}$ buckets of $2^{n/3}$ elements

Computing the sieve

- Count buckets for $\mathcal{A}$ and $\mathcal{B}$
  - $C_{\mathcal{X}}[i] = \left|\left\{x \in \mathcal{X} \mid T(x) = i\right\}\right|$



Computing the sieve

- Count buckets for $\mathcal{A}$ and $\mathcal{B}$
  - $C_{\mathcal{X}}[i] = \left|\left\{x \in \mathcal{X} \;\middle|\; T(x) = i\right\}\right|$
  - $$\begin{aligned} C_{\mathcal{S}}[i] &= |\{(a,b) \in \mathcal{A} \times \mathcal{B} \mid T(a \oplus b) = i\}| \\ &= \sum_{a \in \mathcal{A}} |\{b \in \mathcal{B} \mid T(a \oplus b) = i\}| \\ &= \sum_{a \in \mathcal{A}} C_{\mathcal{B}}[i \oplus T(a)] \\ &= \sum_{j \in \{0,1\}^{n-t}} C_{\mathcal{A}}[j] \cdot C_{\mathcal{B}}[i \oplus j] \end{aligned}$$

Computing the sieve

- Count buckets for $\mathcal{A}$ and $\mathcal{B}$
  - $C_{\mathcal{X}}[i] = \left|\left\{x \in \mathcal{X} \;\middle|\; T(x) = i\right\}\right|$
  - $$\begin{aligned} C_{\mathcal{S}}[i] &= |\{(a,b) \in \mathcal{A} \times \mathcal{B} \mid T(a \oplus b) = i\}| \\ &= \sum_{a \in \mathcal{A}} |\{b \in \mathcal{B} \mid T(a \oplus b) = i\}| \\ &= \sum_{a \in \mathcal{A}} C_{\mathcal{B}}[i \oplus T(a)] \\ &= \sum_{j \in \{0,1\}^{n-t}} C_{\mathcal{A}}[j] \cdot C_{\mathcal{B}}[i \oplus j] \end{aligned}$$
- Discrete convolution can be computed efficiently with the Fast Walsh-Hadamard transform!
  - Complexity: $\tilde{\mathcal{O}}(2^{2n/3})$ for arbitrary $\mathcal{S}$

Then we hope that S is in the bucket with lowest counter:

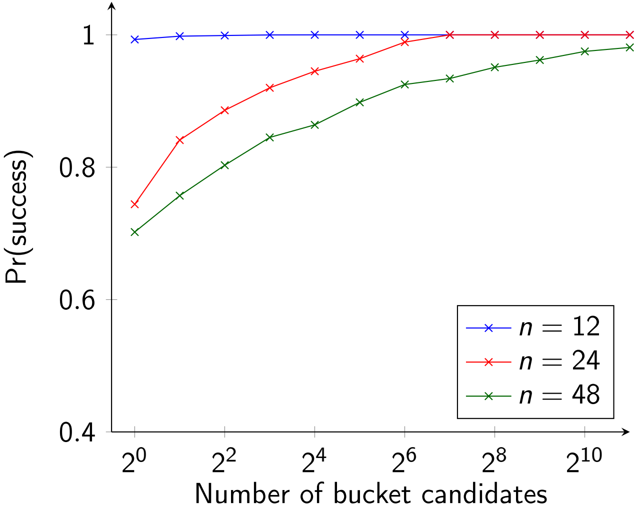
$$T(S) \stackrel{?}{=} \operatorname{argmin} C_{\mathcal{S}}[i]$$

And we can finish with Known-prefix Sieving to recover the rest.

In fact, we can check several candidates and simply hope it is in one of buckets with low counter. The more data, the less bucket candidates we need to try.

Simulation

Figure: Results for $\sqrt{n}2^{2n/3}$ data; counting over $2n/3$ bits.



Missing difference problem algorithms

Algorithms for the missing difference problem

Simple Sieving	Complexity $\tilde{\mathcal{O}}(2^n)$	[McGrew]
Searching	Complexity $\tilde{\mathcal{O}}(2^{n/2}\sqrt{\lvert\mathcal{S}\rvert})$	[McGrew]
Known-prefix Sieving	Complexity $\tilde{\mathcal{O}}(2^{n/2} + 2^{\dim\langle\mathcal{S}\rangle})$	
Fast Convolution Sieving	Complexity $\tilde{\mathcal{O}}(2^{2n/3})$	

Missing difference problem algorithms

Algorithms for the missing difference problem

Simple Sieving	Complexity $\tilde{\mathcal{O}}(2^n)$	[McGrew]
Searching	Complexity $\tilde{\mathcal{O}}(2^{n/2}\sqrt{\lvert\mathcal{S}\rvert})$	[McGrew]
Known-prefix Sieving	Complexity $\tilde{\mathcal{O}}(2^{n/2} + 2^{\dim\langle\mathcal{S}\rangle})$	
Fast Convolution Sieving	Complexity $\tilde{\mathcal{O}}(2^{2n/3})$	

- Improved algorithm if $\mathcal{S}$ is a linear subspace
  - In particular still near optimal when $\dim\langle\mathcal{S}\rangle = n/2$

Missing difference problem algorithms

Algorithms for the missing difference problem

Simple Sieving	Complexity $\tilde{\mathcal{O}}(2^n)$	[McGrew]
Searching	Complexity $\tilde{\mathcal{O}}(2^{n/2}\sqrt{\lvert\mathcal{S}\rvert})$	[McGrew]
Known-prefix Sieving	Complexity $\tilde{\mathcal{O}}(2^{n/2} + 2^{\dim\langle\mathcal{S}\rangle})$	
Fast Convolution Sieving	Complexity $\tilde{\mathcal{O}}(2^{2n/3})$	

- Improved algorithm if $\mathcal{S}$ is a linear subspace
  - In particular still near optimal when $\dim\langle\mathcal{S}\rangle = n/2$
- Improved algorithm for arbitrary $\mathcal{S}$ at the cost of data
  - First algorithm with complexity below 2^n in that case

Back to Cryptanalysis

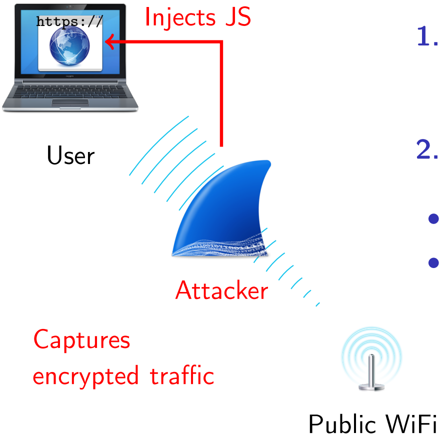
New Tools, New Attacks

Known-prefix $\rightarrow$ plaintext recovery on CTR mode

Fast Convolution $\rightarrow$ forgery on GMAC and Poly1305

First, let's look at a practical setting that gives enough power to the attacker to fully describe an attack.

BEAST Attack Setting [Duong & Rizzo 2011]



- Attacker has access to the network (*eg.* public WiFi)

1. Attacker uses JS to generate traffic
   - Tricks victim to malicious site
   - JS makes *cross-origin* requests
2. Attacker captures encrypted data

- Chosen plaintext attack
- Chosen-Prefix Secret-Suffix model $M \to \mathcal{E}(M \| S)$ [Hoang &al., Crypto'15]

Application to CTR (CPSS queries)

- Plaintext recovery using the known-prefix sieving algorithm
- Two kind of queries; half-block and full-block headers:

Q_1	H_1	S_1	S_2	S_3	S_4	
Q_2	H_1	H_2	S_1	S_2	S_3	S_4

1. Recover S_1 using the first block of each query:

$$\left.\begin{array}{l} \mathcal{A} = \{\mathcal{E}(H_1 \| H_2)\} \\ \mathcal{B} = \{\mathcal{E}(H_1 \| S_1)\} \end{array}\right\} \to \text{Missing difference:} \quad 0 \| (S_1 \oplus H_2).$$

Application to CTR (CPSS queries)

- Plaintext recovery using the known-prefix sieving algorithm
- Two kind of queries; half-block and full-block headers:

Q_1	H_1	S_1	S_2	S_3	S_4	
Q_2	H_1	H_2	S_1	S_2	S_3	S_4

1. Recover S_1 using the first block of each query:
$$\left.\begin{array}{l}\mathcal{A} = \{\mathcal{E}(H_1 \| H_2)\} \\ \mathcal{B} = \{\mathcal{E}(H_1 \| S_1)\}\end{array}\right\} \rightarrow \text{Missing difference:} \quad 0 \| (S_1 \oplus H_2).$$
2. When S_1 is known, recover S_2, with Q_2 queries:
$$\left.\begin{array}{l}\mathcal{A} = \{\mathcal{E}(H_1 \| H_2)\} \\ \mathcal{B} = \{\mathcal{E}(S_1 \| S_2)\}\end{array}\right\} \rightarrow \text{Missing difference:} \quad (S_1 \oplus H_1) \| (S_2 \oplus H_2).$$

Application to CTR (CPSS queries)

- Plaintext recovery using the known-prefix sieving algorithm
- Two kind of queries; half-block and full-block headers:

Q_1	H_1	S_1	S_2	S_3	S_4	
Q_2	H_1	H_2	S_1	S_2	S_3	S_4

1. Recover S_1 using the first block of each query:
$$\left.\begin{array}{l}\mathcal{A} = \{\mathcal{E}(H_1 \| H_2)\} \\ \mathcal{B} = \{\mathcal{E}(H_1 \| S_1)\}\end{array}\right\} \to \text{Missing difference:} \quad 0 \| (S_1 \oplus H_2).$$
2. When S_1 is known, recover S_2, with Q_2 queries:
$$\left.\begin{array}{l}\mathcal{A} = \{\mathcal{E}(H_1 \| H_2)\} \\ \mathcal{B} = \{\mathcal{E}(S_1 \| S_2)\}\end{array}\right\} \to \text{Missing difference:} \ (S_1 \oplus H_1) \| (S_2 \oplus H_2).$$
3. When S_2 is known, recover S_3:
$$\left.\begin{array}{l}\mathcal{A} = \{\mathcal{E}(H_1 \| H_2)\} \\ \mathcal{B} = \{\mathcal{E}(S_2 \| S_3)\}\end{array}\right\} \to \text{Missing difference:} \ (S_2 \oplus H_1) \| (S_3 \oplus H_2).$$
4. ...

Application to CTR (CPSS queries)

Remarks on this attack:

- We perform the Known-prefix sieving twice per block of secret.
- We reuse queries so we don't need additional queries to uncover additional blocks of secret.
    - Once you gathered enough queries to recover S_1 and S_2 it is probably enough to recover all of the secret.

Application to CTR (CPSS queries)

Remarks on this attack:

- We perform the Known-prefix sieving twice per block of secret.
- We reuse queries so we don't need additional queries to uncover additional blocks of secret.
    - Once you gathered enough queries to recover S_1 and S_2 it is probably enough to recover all of the secret.

Full Asymptotic Complexity

Queries	$\mathcal{O}(\sqrt{n} \cdot 2^{n/2})$
Memory	$\mathcal{O}(\sqrt{n} \cdot 2^{n/2})$
Time	$\mathcal{O}(n \cdot 2^{n/2})$

Wegman-Carter Authentication Modes

- Wegman-Carter: build a MAC from a universal hash function and a PRF

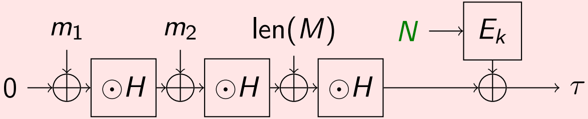
$$\mathrm{WC}(N, M) = H_{k_1}(M) \oplus F_{k_2}(N).$$

$$\mathbf{Adv}^{\mathrm{MAC}}_{\mathrm{WC}[H,F]} \leq \mathbf{Adv}^{\mathrm{PRF}}_{F} + \varepsilon + 2^{-n}$$

- Wegman-Carter-Shoup: use a block cipher as a PRF

$$\mathrm{WCS}(N, M) = H_{k_1}(M) \oplus E_{k_2}(N),$$

Example: Polynomial-based hashing (GMAC, Poly1305-AES)



Application to GMAC

Authentication of one block A of authenticated data in a given Galois field:

$$\mathrm{MAC}(N, A) = A \cdot H^2 \oplus H \oplus E_k(N)$$

with N a never repeating nonce, H the hash key.

Application to GMAC

Authentication of one block A of authenticated data in a given Galois field:

$$\mathrm{MAC}(N, A) = A \cdot H^2 \oplus H \oplus E_k(N)$$

with N a never repeating nonce, H the hash key.
Collect many signatures for A and A', then $\forall i \neq j$:

$$\begin{aligned}\mathrm{MAC}(i, A) \oplus \mathrm{MAC}(j, A') &\neq A \cdot H^2 \oplus H \oplus A' \cdot H^2 \oplus H \\ &\neq (A \oplus A') \cdot H^2\end{aligned}$$

Application to GMAC

Authentication of one block A of authenticated data in a given Galois field:

$$\mathrm{MAC}(N, A) = A \cdot H^2 \oplus H \oplus E_k(N)$$

with N a never repeating nonce, H the hash key.
Collect many signatures for A and A', then $\forall i \neq j$:

$$\begin{aligned}\mathrm{MAC}(i, A) \oplus \mathrm{MAC}(j, A') &\neq A \cdot H^2 \oplus H \oplus A' \cdot H^2 \oplus H \\ &\neq (A \oplus A') \cdot H^2\end{aligned}$$

- Solve the missing difference problem.
- Invert $A \oplus A'$, get H^2.
- Find the square root, get H, the hash key!

Key recovery as a missing difference problem

- Fix two messages $M \neq M'$, capture MACs
  - $a_i = \mathrm{MAC}(i, M) = H_{K_1}(M) \oplus K_i$
  - $b_j = \mathrm{MAC}(j, M') = H_{K_1}(M') \oplus K_j$
  - $a_i \oplus b_j \neq H_{K_1}(M) \oplus H_{K_1}(M')$
- For polynomial hashing, easy to recover universal hash key from $H_{K_1}(M) \oplus H_{K_1}(M')$

Key recovery as a missing difference problem

- Fix two messages $M \neq M'$, capture MACs
  - $a_i = \mathrm{MAC}(i, M) = H_{K_1}(M) \oplus K_i$
  - $b_j = \mathrm{MAC}(j, M') = H_{K_1}(M') \oplus K_j$
  - $a_i \oplus b_j \neq H_{K_1}(M) \oplus H_{K_1}(M')$
- For polynomial hashing, easy to recover universal hash key from $H_{K_1}(M) \oplus H_{K_1}(M')$
- Sieving algorithm recovers $H(M) \oplus H(M')$ with $\tilde{\mathcal{O}}(2^{n/2})$ queries and $\tilde{\mathcal{O}}(2^n)$ computations
  - Independently done in another Eurocrypt paper!

Optimal Forgeries Against Polynomial-Based MACs and GCM
Atul Luykx, Bart Preneel [Eurocrypt '18]

Key recovery as a missing difference problem

- Fix two messages $M \neq M'$, capture MACs
  - $a_i = \mathrm{MAC}(i, M) = H_{K_1}(M) \oplus K_i$
  - $b_j = \mathrm{MAC}(j, M') = H_{K_1}(M') \oplus K_j$
  - $a_i \oplus b_j \neq H_{K_1}(M) \oplus H_{K_1}(M')$
- For polynomial hashing, easy to recover universal hash key from $H_{K_1}(M) \oplus H_{K_1}(M')$
- Sieving algorithm recovers $H(M) \oplus H(M')$ with $\tilde{\mathcal{O}}(2^{n/2})$ queries and $\tilde{\mathcal{O}}(2^n)$ computations
  - Independently done in another Eurocrypt paper!

    Optimal Forgeries Against Polynomial-Based MACs and GCM
    Atul Luykx, Bart Preneel [Eurocrypt '18]
- Fast convolution sieving recovers $H(M) \oplus H(M')$ with $\tilde{\mathcal{O}}(2^{2n/3})$ queries and computations
  - First universal forgery attack with less than 2^n operations

Impacts

How practical can be the plaintext recovery attack on CTR ?

- Mostly used with AES, famous 128-bit block cipher, as part of GCM. 90% of Firefox HTTPS traffic uses AES-GCM.
  - Requires 128×2^{64} bits $=$ 256 exbibytes over one session
  - 2016 global IP traffic is 82.3 exbibytes per month [Cisco]

Impacts

How practical can be the plaintext recovery attack on CTR ?

- Mostly used with AES, famous 128-bit block cipher, as part of GCM. 90% of Firefox HTTPS traffic uses AES-GCM.
  - Requires 128×2^{64} bits = 256 exbibytes over one session
  - 2016 global IP traffic is 82.3 exbibytes per month [Cisco]
- SSHv2 implements CTR with 3DES, a 64-bit block cipher.
  - Requires 64×2^{32} bits = 32 gibibytes
  - Quickly attainable with modern internet speed

Impacts

How practical can be the plaintext recovery attack on CTR ?

- Mostly used with AES, famous 128-bit block cipher, as part of GCM. 90% of Firefox HTTPS traffic uses AES-GCM.
  - Requires 128×2^{64} bits = 256 exbibytes over one session
  - 2016 global IP traffic is 82.3 exbibytes per month [Cisco]
- SSHv2 implements CTR with 3DES, a 64-bit block cipher.
  - Requires 64×2^{32} bits = 32 gibibytes
  - Quickly attainable with modern internet speed

Sweet32 attack

Attack in the BEAST setting with birthday bound complexity already shown to be a threat over the web in previous work by Bhargavan and Leurent.
This is the **Sweet32** attack on CBC mode, more commonly used with 64-bit block ciphers.

Counter-measures

1. Use AES, or any good 128-bit block cipher.
   - Make n big enough so that $2^{n/2}$ is impractical.
   - Most obvious choice for most new implementations.

Counter-measures

1. Use AES, or any good 128-bit block cipher.
   - Make n big enough so that $2^{n/2}$ is impractical.
   - Most obvious choice for most new implementations.

2. Forget block ciphers, use a PRF.
   - CTR is perfectly secure as long as we use a good PRF.
   - Dedicated PRF are rare but many solutions exist (XoP).

Counter-measures

1. Use AES, or any good 128-bit block cipher.
   - Make n big enough so that $2^{n/2}$ is impractical.
   - Most obvious choice for most new implementations.

2. Forget block ciphers, use a PRF.
   - CTR is perfectly secure as long as we use a good PRF.
   - Dedicated PRF are rare but many solutions exist (XoP).

3. Forget CTR, use advanced Beyond Birthday Bound schemes.
   - They have a proof with better security bounds.
   - CENC is a BBB scheme derived from CTR. [Iwata, FSE'06]

Counter-measures

1. Use AES, or any good 128-bit block cipher.
   - Make n big enough so that $2^{n/2}$ is impractical.
   - Most obvious choice for most new implementations.
2. Forget block ciphers, use a PRF.
   - CTR is perfectly secure as long as we use a good PRF.
   - Dedicated PRF are rare but many solutions exist (XoP).
3. Forget CTR, use advanced Beyond Birthday Bound schemes.
   - They have a proof with better security bounds.
   - CENC is a BBB scheme derived from CTR. [Iwata, FSE'06]
4. Simply rekey frequently.
   - Rekeying way before $2^{n/2}$ blocks efficiently prevents the attack.
   - Maybe the easiest hotfix.

Conclusion

Case	Previous	This work	Improved attacks
$\mathcal{S}$ affine subspace of dim $n/2$.	$\tilde{\mathcal{O}}(2^{3n/4})$	$\tilde{\mathcal{O}}(2^{n/2})$	CTR plaintext recovery.
No prior info on S. *ie.* $\lvert\mathcal{S}\rvert = 2^n$.	$\tilde{\mathcal{O}}(2^{n})$	$\tilde{\mathcal{O}}(2^{2n/3})$	GMAC, Poly1305 universal forgery.

Conclusion

Case	Previous	This work	Improved attacks
$\mathcal{S}$ affine subspace of dim $n/2$.	$\tilde{\mathcal{O}}(2^{3n/4})$	$\tilde{\mathcal{O}}(2^{n/2})$	CTR plaintext recovery.
No prior info on S. *ie.* $\lvert\mathcal{S}\rvert = 2^n$.	$\tilde{\mathcal{O}}(2^{n})$	$\tilde{\mathcal{O}}(2^{2n/3})$	GMAC, Poly1305 universal forgery.

Especially when $n = 64$ bits, main take away :

- CTR mode not more secure than CBC (Sweet32).
- Frequent rekeying away from birthday bound will prevent these attacks.

Fast Walsh-Hadamard transform

We need an efficient algorithm to compute the multiplication of a Hadamard matrix H_m by a vector of size 2^m in $\mathcal{O}(m \cdot 2^m)$.

$$H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_m = \frac{1}{2^{m/2}} H^{\otimes m}$$

That is the fast Walsh-Hadamard transform (FWHT), akin to a fast Fourier transform.

Fast XOR-counting

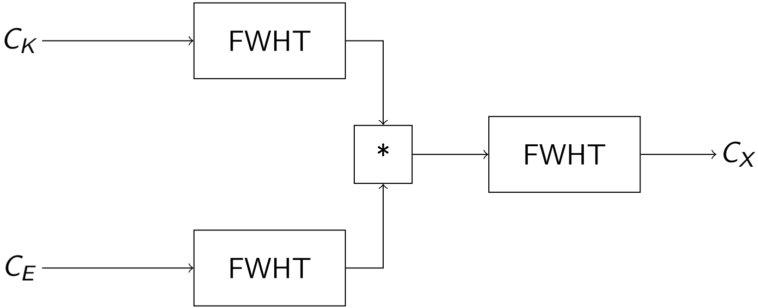


Figure: Fast XOR-counting algorithm

Fast XOR-counting

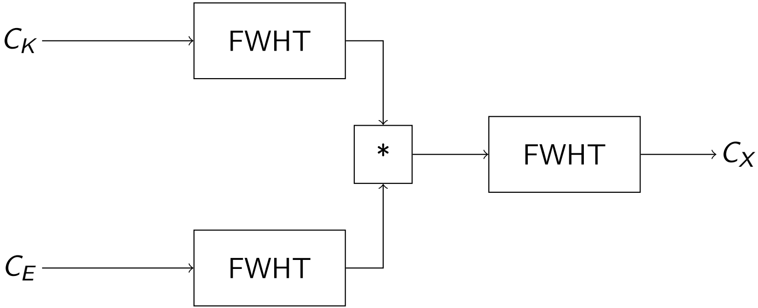


Figure: Fast XOR-counting algorithm

Note that $\mathrm{FWHT}^{-1} = \mathrm{FWHT}$.
We hope that :

$$S_{2n/3} \stackrel{?}{=} \operatorname*{argmin}_i C_X[i]$$

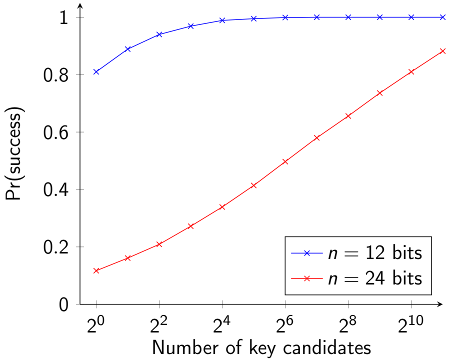
Fast XOR-counting

For an $\Omega(1)$ probability of success on the first trial assuming independence of the counters (/!\ False as $\sum C_X = |\mathcal{K} \times \mathcal{E}|$.) :

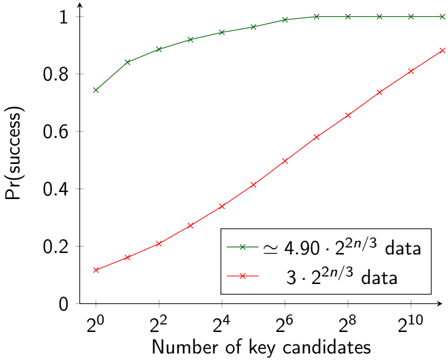
Complexity

$\mathcal{O}(\sqrt{n} \cdot 2^{2n/3})$	*queries*
$\mathcal{O}(n \cdot 2^{2n/3}) + \mathcal{O}(n\sqrt{n} \cdot 2^{n/2})$	*bits memory (counters + sieving)*
$\mathcal{O}(n \cdot 2^{2n/3}) + \mathcal{O}(n\sqrt{n} \cdot 2^{n/2})$	*computations (FWHT + sieving)*

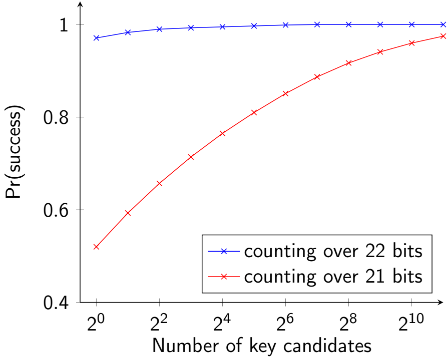
(a) Results for lists size of $3 \cdot 2^{2n/3}$



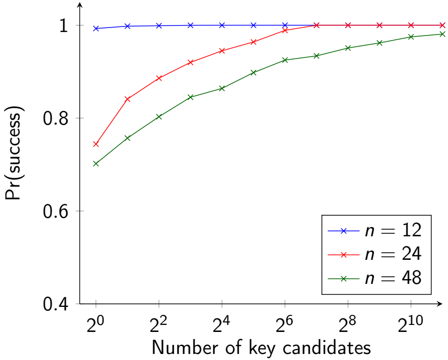
(b) Results for $n = 24$ bits



(a) Results for $n = 32$ bits; $\sqrt{n}2^{2n/3} \simeq 5.66 \cdot 2^{2n/3}$ data



(b) Results for $\sqrt{n}2^{2n/3}$ data; counting over $2n/3$ bits



Poly1305

For a key r, some nonce N and message M of length q the Poly1305's MAC is defined as:

$$T(M,N) = ((c_1 r^q + c_2 r^{q-1} + ... + c_q r) \mod 2^{130} - 5) + E_k(N) \mod 2^{128}$$

Poly1305

For a key r, some nonce N and message M of length q the Poly1305's MAC is defined as:

$$T(M,N) = ((c_1 r^q + c_2 r^{q-1} + ... + c_q r) \mod 2^{130} - 5) + E_k(N) \mod 2^{128}$$

Then for two messages M, M' the missing difference will be :

$$((c_1 - c_1') r^q + (c_2 - c_2') r^{q-1} + ... + (c_q - c_q') r) \mod 2^{130} - 5 \mod 2^{128}$$

Poly1305

For a key r, some nonce N and message M of length q the Poly1305's MAC is defined as:

$$T(M,N) = ((c_1 r^q + c_2 r^{q-1} + ... + c_q r) \mod 2^{130} - 5) + E_k(N) \mod 2^{128}$$

Then for two messages M, M' the missing difference will be :

$$((c_1 - c_1')r^q + (c_2 - c_2')r^{q-1} + ... + (c_q - c_q')r) \mod 2^{130} - 5 \mod 2^{128}$$

Choose M and M' so that $(c_q - c_q') = 1$, $(c_i - c_i') = 0$ and the missing difference will be r as $r < 2^{124}$ by construction. This is the hash key!

Poly1305

For a key r, some nonce N and message M of length q the Poly1305's MAC is defined as:

$$T(M,N) = ((c_1 r^q + c_2 r^{q-1} + \ldots + c_q r) \mod 2^{130} - 5) + E_k(N) \mod 2^{128}$$

Then for two messages M, M' the missing difference will be :

$$((c_1 - c_1')r^q + (c_2 - c_2')r^{q-1} + \ldots + (c_q - c_q')r) \mod 2^{130} - 5 \mod 2^{128}$$

Choose M and M' so that $(c_q - c_q') = 1$, $(c_i - c_i') = 0$ and the missing difference will be r as $r < 2^{124}$ by construction. This is the hash key!

Note : As we play with modular addition and not xor operation we have to compute a cyclic convolution using fast Fourier transform instead of Walsh-Hadamard.