

Breaking and Repairing GCM Security Proofs (from CRYPTO 2012)

Tetsu Iwata, Nagoya University

Keisuke Ohashi, Nagoya University

Kazuhiko Minematsu, NEC Corporation

ASK 2012

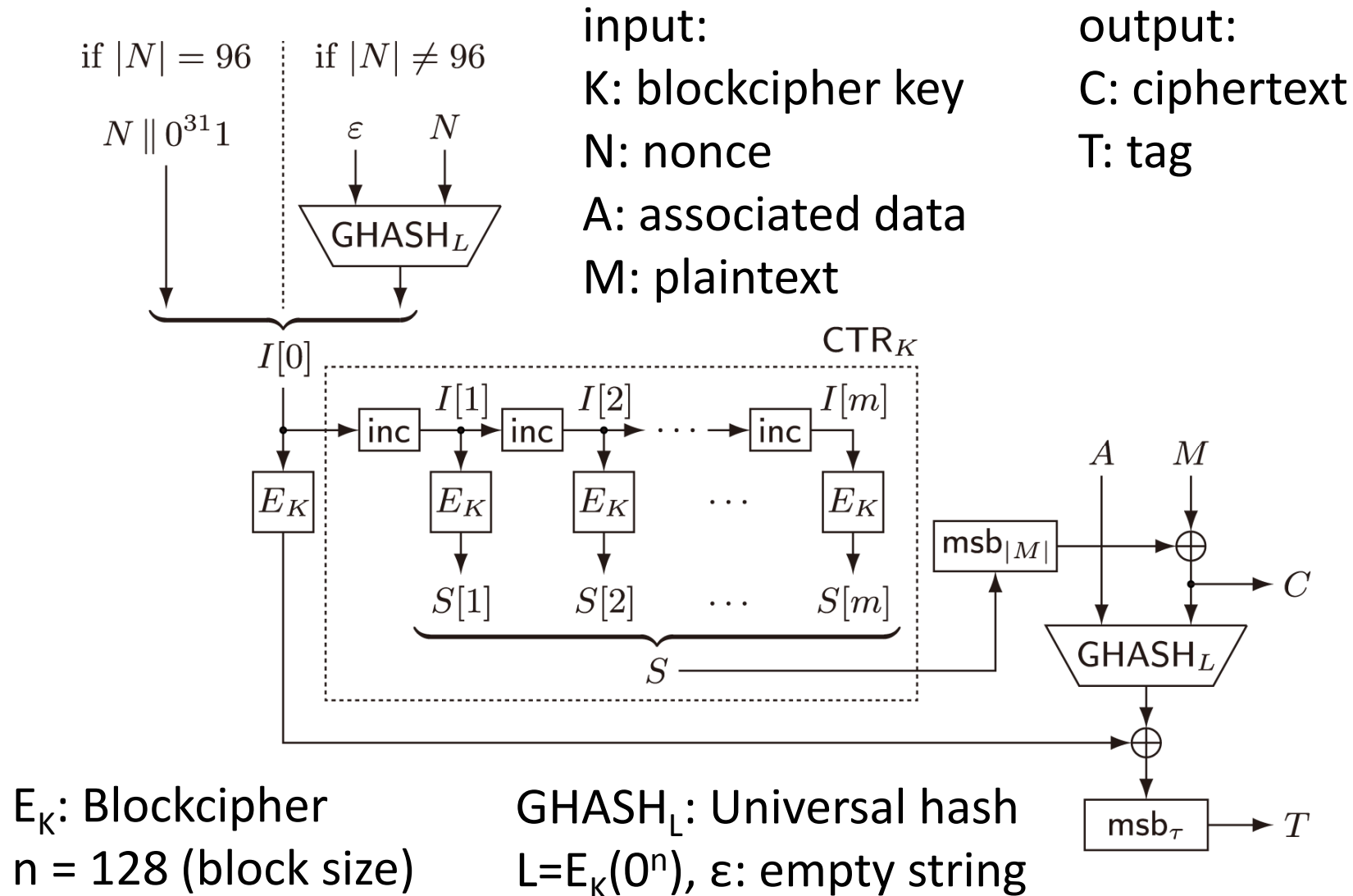
August 27, 2012, Nagoya, Japan

GCM

- Galois/Counter Mode
- authenticated encryption mode of 128-bit blockciphers
- designed by McGrew and Viega in 2004 [MV04]
- selected as the NIST recommended authenticated encryption mode in 2007
- widely used in practice
 - ISO/IEC 19772, IEEE P1619.1, NSA Suite B, IETF IPsec, SSH, SSL,...

[MV04] David A. McGrew and John Viega: The Security and Performance of the Galois/Counter Mode (GCM) of Operation. Cryptology ePrint Archive: Report 2004/193 (full version of INDOCRYPT 2004)

Encryption Algorithm of GCM



Provable Security Results

- The designers proved the security of GCM [MV04]
- analyzed privacy and authenticity against chosen ciphertext attacks

- Privacy bound:

$$\begin{aligned} - \text{Adv}_{\text{GCM}[\text{Perm}(n), \tau]}^{\text{priv}}(\mathcal{A}) \leq & \frac{0.5(\sigma/n + 2q)^2}{2^n} \\ & + \frac{2q(\sigma/n + 2q)\lceil \ell_N/n + 1 \rceil}{2^n} + \frac{q\lceil \ell/n + 1 \rceil}{2^\tau} \end{aligned}$$

- Ciphertexts of GCM are indistinguishable from random strings

Provable Security Results

- Authenticity bound:

$$\begin{aligned} - \text{Adv}_{\text{GCM}[\text{Perm}(n), \tau]}^{\text{auth}}(\mathcal{A}) \leq & \frac{0.5(\sigma/n + 2q)^2}{2^n} \\ & + \frac{2q(\sigma/n + 2q + 1)\lceil \ell_N/n + 1 \rceil}{2^n} + \frac{q\lceil \ell/n + 1 \rceil}{2^\tau} \end{aligned}$$

- GCM is unforgeable

Previous Security Analyses

- [Ferguson '05]
 - forgery attacks when the tag is short
- [Joux '06]
 - key recovery attacks on GCM (nonce reuse), forgery attacks on the draft NIST version of GCM
- [Handschuh, Preneel '08]
 - a weak key, forgery attacks
- [Saarinen '12]
 - many weak keys, forgery attacks

Previous Security Analyses

- It is widely considered that the provable security results of GCM are sound
 - in the sense that these attacks do not contradict the claimed security bounds, and that no flaw in the proofs has been identified
 - show the tightness of the security bounds
 - outside the security model (e.g., nonce reuse)

Equation Over $GF(2^{128})$

- defined by the irreducible polynomial $p(x) = 1+x+x^2+x^7+x^{128}$ (used in GCM)
- the multiplicative identity element is $0x80...0$ ($10...0$ in binary)

$$U \cdot L^2 \oplus V \cdot L \oplus 0x0...01 = U' \cdot L^2 \oplus V \cdot L$$

$U = 0x00000000\ 00000000\ 02000000\ 00000000$ (128 bits)

$U' = 0x00000000\ 00000000\ 06000000\ 00000000$ (128 bits)

$V = 0x00000000\ 00000000\ 00000000\ 00000048$ (128 bits)

- How many solutions (in L) do we have?

Equation Over $GF(2^{128})$

- defined by the irreducible polynomial $p(x) = 1+x+x^2+x^7+x^{128}$ (used in GCM)
- the multiplicative identity element is $0x80...0$ ($10...0$ in binary)

$$U \cdot L^2 \oplus V \cdot L \oplus 0x0...01 = U' \cdot L^2 \oplus V \cdot L$$

$U = 0x00000000\ 00000000\ 02000000\ 00000000$ (128 bits)

$U' = 0x00000000\ 00000000\ 06000000\ 00000000$ (128 bits)

$V = 0x00000000\ 00000000\ 00000000\ 00000048$ (128 bits)

- How many solutions (in L) do we have?
 - at most 2 solutions (actually one solution)

Equation Over $GF(2^{128})$

$$U \cdot L^2 \oplus V \cdot L \oplus 0x0...01 = U' \cdot L^2 \oplus V \cdot L$$

$$U \cdot L^2 \oplus 0x0...01 = U' \cdot L^2$$

$$(U \oplus U') \cdot L^2 = 0x0...01$$

$$L^2 = (U \oplus U')^{-1}$$

$$L = (U \oplus U')^{-2^{127}}$$

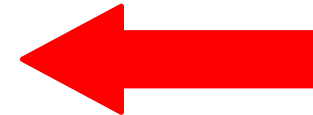
- How many solutions (in L) do we have?
 - at most 2 solutions (actually one solution)

Increment Function in GCM

- $\text{inc}(X \parallel Y) = X \parallel (Y+1 \bmod 2^{32})$
 - $|X| = 96, |Y| = 32$
 - $\text{inc}(0x0\dots01) = 0x0\dots02$

$$U \cdot L^2 \oplus V \cdot L \oplus \underline{0x0\dots01} = U' \cdot L^2 \oplus V \cdot L$$

$$\text{inc}(U \cdot L^2 \oplus V \cdot L) = U' \cdot L^2 \oplus V \cdot L$$



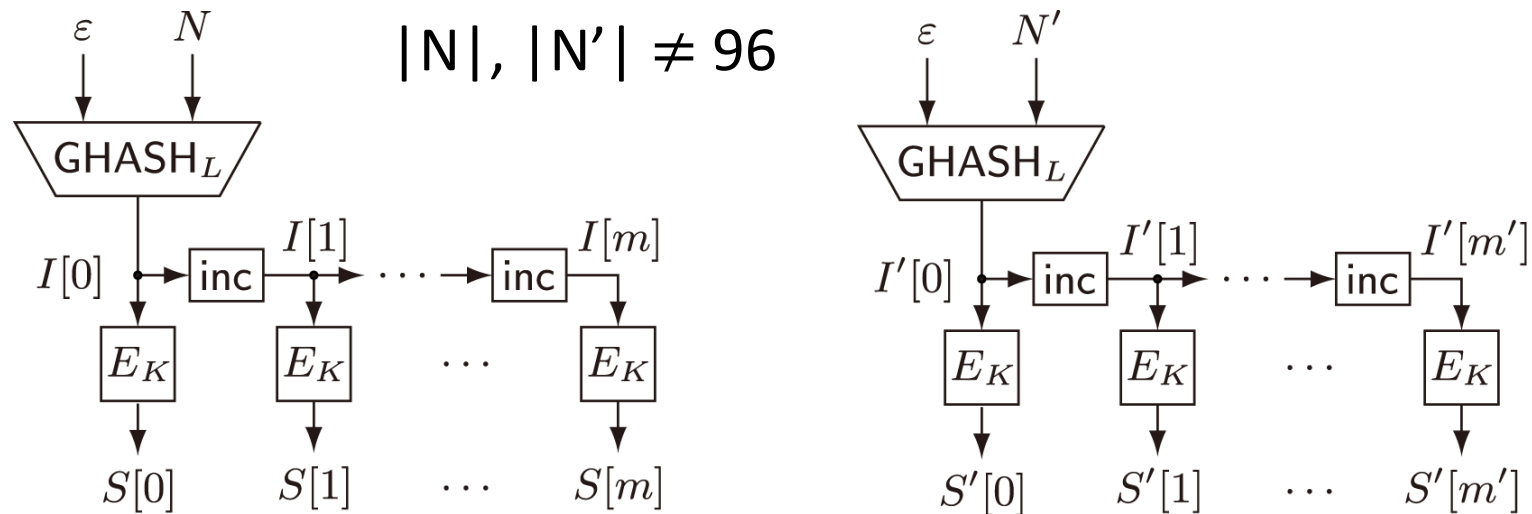
- How many solutions (in L) do we have?
 - Note: LHS may not be a degree 2 polynomial over $\text{GF}(2^{128})$

List of Solutions

0x7f6db6d2db6db6db6db6db6492492492, 0x7f6db6dadb6db6db6db6db6492492492,
0x81b6db776db6db6db6db6dadb6db6db6, 0x81b6db676db6db6db6db6dadb6db6db6,
0xbe00003c0000000000000003fffffffff, 0xbe00001c0000000000000003fffffffff,
0xc16db6aadb6db6db6db6db6db1b6db6db6d, 0xc16db6eadb6db6db6db6db6db1b6db6db6d,
0x3fb6db876db6db6db6db6d5249249249, 0x3fb6db076db6db6db6db6d5249249249,
0x000001dc0000000000000001c000000000, 0x000000dc0000000000000001c000000000,
0x7f6db56adb6db6db6db6db6d8e492492492, 0x7f6db76adb6db6db6db6d8e492492492,
0x81b6dc076db6db6db6db6aadb6db6db6, 0x81b6d8076db6db6db6db6aadb6db6db6,
0xbe000edc000000000000000e3fffffffff, 0xbe0006dc000000000000000e3fffffffff,
0xc16dab6adb6db6db6db6c71b6db6db6d, 0xc16dbb6adb6db6db6db6c71b6db6db6d,
0x3fb6e0076db6db6db6db6555249249249, 0x3fb6c0076db6db6db6db6555249249249,
0x000076dc000000000000071c0000000000, 0x000036dc000000000000071c0000000000,
0x7f6d5b6adb6db6db6db638e492492492, 0x7f6ddb6adb6db6db6db638e492492492,
0x81b700076db6db6db6daaaadb6db6db6, 0x81b600076db6db6db6daaaadb6db6db6,
0xbe03b6dc0000000000038e3fffffffff, 0xbe01b6dc0000000000038e3fffffffff,
0xc16adb6adb6db6db6db6c71b6db6db6d, 0x0000000400000000000000000000000000000000

- Answer: 32 solutions

Counter Collision



- A counter collision is a bad event: $I[1] = I'[1], I[2] = I'[2], \dots$
 - xor of two ciphertexts = xor of two plaintexts
 - the information about plaintexts is leaked
- We need to show that $\Pr_L[\text{Coll}_L(r, N, N')]]$ is small
 - $\text{Coll}_L(r, N, N') : \text{inc}^r(\text{GHASH}_L(\epsilon, N)) = \text{GHASH}_L(\epsilon, N')$

GHASH_L(ε, N)

- universal hash function
- $N \parallel 0 \dots 0 \parallel |N|_n = (X[1], \dots, X[x])$
- $\text{GHASH}_L(\epsilon, N) = X[1] \cdot L^x \oplus X[2] \cdot L^{x-1} \oplus \dots \oplus X[x] \cdot L$
- $N = 0x00000000\ 00000000\ 02$ (72 bits)
- $\text{GHASH}_L(\epsilon, N)$

$$= 0x00000000\ 00000000\ 02000000\ 00000000 \cdot L^2$$

$$\oplus 0x00000000\ 00000000\ 00000000\ 00000048 \cdot L$$

$$= U \cdot L^2 \oplus V \cdot L$$
- $N' = 0x00000000\ 00000000\ 06$ (72 bits)
- $\text{GHASH}_L(\epsilon, N') = U' \cdot L^2 \oplus V \cdot L$

$\Pr_L[\text{Coll}_L(r, N, N')]$ Is Small

- [Lemma 3, MV04]

$$\Pr_L[\text{Coll}_L(r, N, N')] \leq \max\{ d, d' \} / 2^n$$

where $d = \deg(\text{GHASH}_L(\epsilon, N))$, $d' = \deg(\text{GHASH}_L(\epsilon, N'))$

- The lemma claims

“ $\text{inc}(U \cdot L^2 \oplus V \cdot L) = U' \cdot L^2 \oplus V \cdot L$ has at most 2 solutions.”

$\Pr_L[\text{Coll}_L(r, N, N')]$ Is Small

- [Lemma 3, MV04]

$$\Pr_L[\text{Coll}_L(r, N, N')] \leq \max\{ d, d' \} / 2^n$$

where $d = \deg(\text{GHASH}_L(\epsilon, N))$, $d' = \deg(\text{GHASH}_L(\epsilon, N'))$

- The lemma claims

“ $\text{inc}(U \cdot L^2 \oplus V \cdot L) = U' \cdot L^2 \oplus V \cdot L$ has at most 2 solutions.”

- [Lemma 3, MV04] is incorrect
 - used in both the privacy proof and the authenticity proof
 - both proofs contain a flaw

More Observation

$$\text{inc}(U \cdot L^2 \oplus V \cdot L) = U' \cdot L^2 \oplus V \cdot L \quad (\text{A})$$

$$\text{inc}^2(U \cdot L^2 \oplus V \cdot L) = U' \cdot L^2 \oplus V \cdot L \quad (\text{B})$$

$$\text{inc}^4(U \cdot L^2 \oplus V \cdot L) = U' \cdot L^2 \oplus V \cdot L \quad (\text{C})$$

$$\text{inc}^0(U \cdot L^2 \oplus V \cdot L) = U' \cdot L^2 \oplus V \cdot L \quad (\text{D})$$

- Number of solutions
 - (A): 32, (B): 31, (C): 30, (D): 1
- 94 solutions are all distinct
- $\Pr_L[(\text{A}) \text{ or } (\text{B}) \text{ or } (\text{C}) \text{ or } (\text{D})] \geq 94/2^{128}$

Distinguishing Attack

- The observation can be translated into a distinguishing attack on $\text{GCM}[\text{Rand}(n), \tau]$: GCM with a random function R (instead of E_K)
 - by simply observing if the event occurs in ciphertexts
 - $\text{Adv}_{\text{GCM}[\text{Rand}(n), \tau]}^{\text{priv}}(A) \geq 94/2^{128}$
- The attack does not contradict the overall privacy bound, but it invalidates a part of it
 - $\text{Adv}_{\text{GCM}[\text{Perm}(n), \tau]}^{\text{priv}}(\mathcal{A}) \leq \frac{0.5(\sigma/n + 2q)^2}{2^n} + \frac{2q(\sigma/n + 2q)\lceil \ell_N/n + 1 \rceil}{2^n} + \frac{q\lceil \ell/n + 1 \rceil}{2^\tau}$
 - The second term: $\text{Adv}_{\text{GCM}[\text{Rand}(n), \tau]}^{\text{priv}}(A) \leq 80/2^{128}$

Remarks

- The attack does not break GCM
 - Our attack does not contradict the overall privacy bound
 - it invalidates only a part of it
 - the attack also invalidates a part of the authenticity bound
- The success probability of the attack is small
 - The practical implication is limited
- The attack does not work if the nonce length is restricted to 96 bits (required or recommended by many standards mainly for efficiency reasons)

Can We Repair the Proofs?

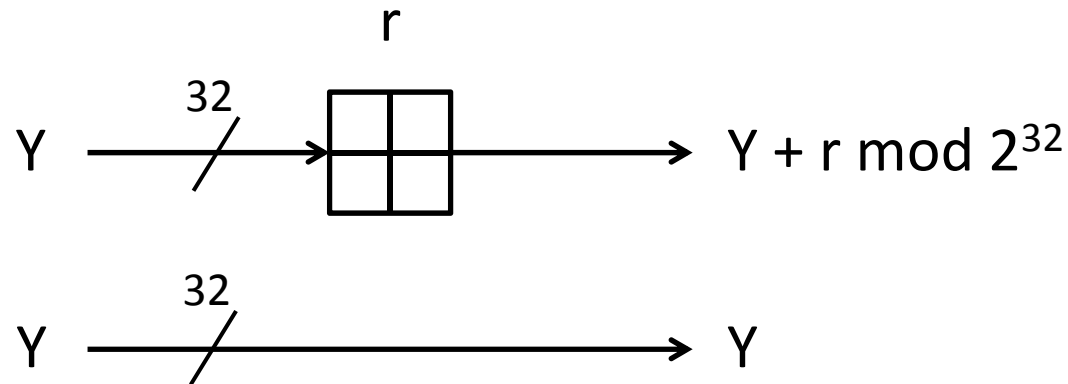
- without modifying the original specification
- $\Pr_L[\text{Coll}_L(r, N, N')] \leq ?$
 - introduce a combinatorial problem
 - relation to the proof
 - approaches to solve the problem
 - new privacy and authenticity bounds

Combinatorial Problem

$$\mathbf{Y}_r = \{ (Y + r \bmod 2^{32}) \oplus Y \mid Y \text{ is in } \{0,1\}^{32} \}$$

$$\alpha_r = \# \mathbf{Y}_r$$

problem: determine $\alpha_{\max} = \max\{ \alpha_r \mid 0 \leq r \leq 2^{32}-1 \}$



α_r = the number of possible non-zero xor differences of $Y + r \bmod 2^{32}$ and Y when Y ranges over $\{0,1\}^{32}$

Combinatorial Problem

$$\mathbf{Y}_r = \{ (Y + r \bmod 2^{32}) \oplus Y \mid Y \text{ is in } \{0,1\}^{32} \}$$

$$\alpha_r = \# \mathbf{Y}_r$$

problem: determine $\alpha_{\max} = \max\{ \alpha_r \mid 0 \leq r \leq 2^{32}-1 \}$

- Toy example: $\mathbf{Y}_1 = \{ (Y + 1 \bmod 2^3) \oplus Y \mid Y \text{ is in } \{0,1\}^3 \}$

$$- (000 + 001) \oplus 000 = 001$$

$$- (001 + 001) \oplus 001 = 011$$

$$- (010 + 001) \oplus 010 = 001$$

$$- (011 + 001) \oplus 011 = 111$$

$$- (100 + 001) \oplus 100 = 001$$

$$- (101 + 001) \oplus 101 = 011$$

$$- (110 + 001) \oplus 110 = 001$$

$$- (111 + 001) \oplus 111 = 111$$

$$\mathbf{Y}_1 = \{ 001, 011, 111 \}$$

$$\alpha_1 = 3$$

Relation to the Proof

- $\text{Coll}_L(r, N, N') : \text{inc}^r(\text{GHASH}_L(\varepsilon, N)) = \text{GHASH}_L(\varepsilon, N')$
- if we can replace $\text{inc}^r(\text{GHASH}_L(\varepsilon, N))$ by $\text{GHASH}_L(\varepsilon, N) \oplus C$, then we can derive the upper bound on $\Pr_L[\text{Coll}_L(r, N, N')]$

$$\text{GHASH}_L(\varepsilon, N) \oplus C = \text{GHASH}_L(\varepsilon, N') \quad (*)$$

- but C depends on r and $\text{GHASH}_L(\varepsilon, N)$
- $\alpha_r = \#\{ (Y + r \bmod 2^{32}) \oplus Y \mid Y \text{ is in } \{0,1\}^{32} \}$ represents the number of possibilities of C
- For each C , we know the number of solutions of $(*)$

Relation to the Proof

- Towards a new version of [Lemma 3, MV04]
- Lemma

For each $0 \leq r \leq 2^{32}-1$

$$\Pr_L[\text{Coll}_L(r, N, N')] \leq \alpha_r \max\{ d, d' \} / 2^n$$

where $d = \deg(\text{GHASH}_L(\varepsilon, N))$, $d' = \deg(\text{GHASH}_L(\varepsilon, N'))$

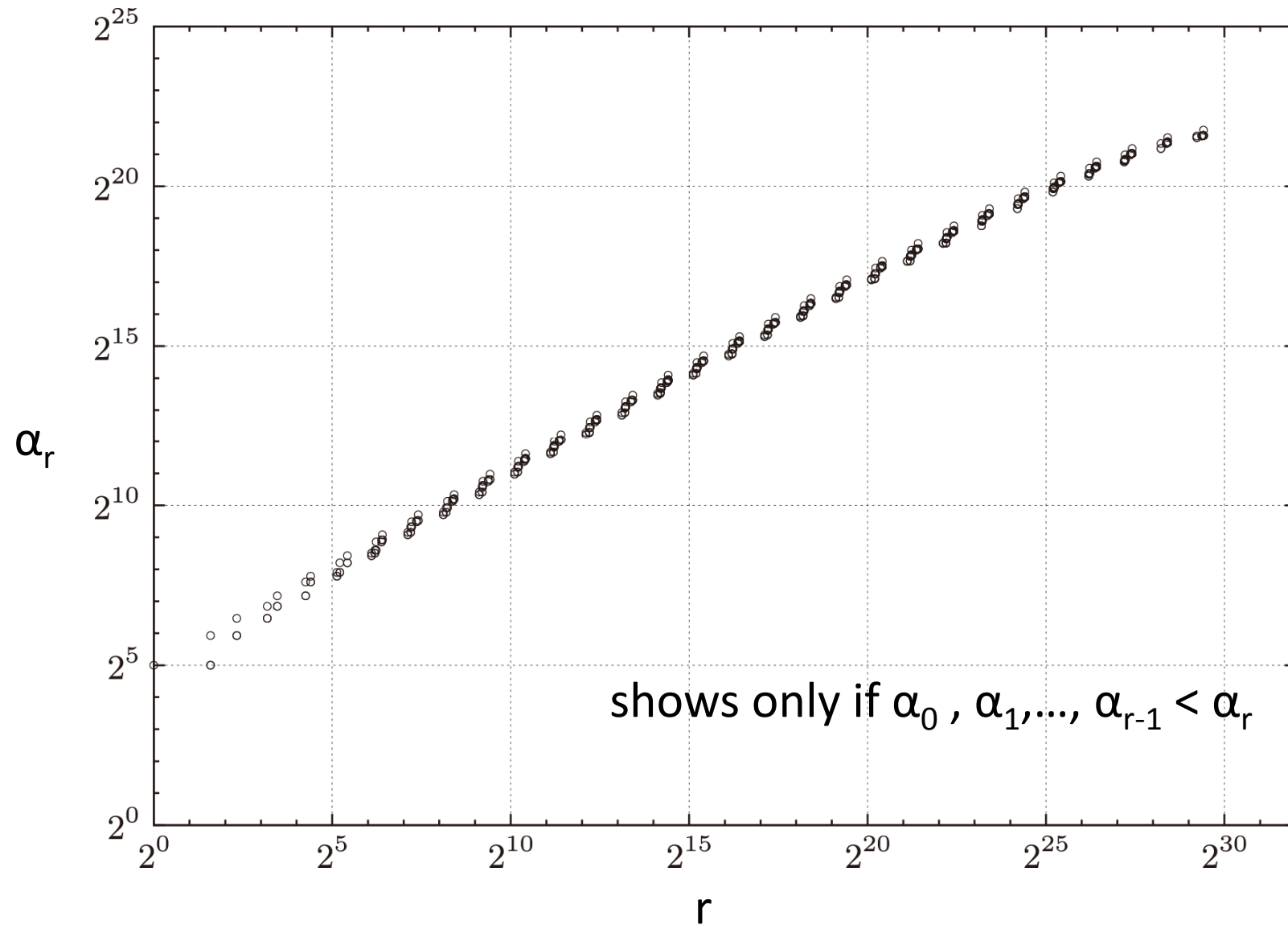
- For any $0 \leq r \leq 2^{32}-1$

$$\Pr_L[\text{Coll}_L(r, N, N')] \leq \alpha_{\max} \max\{ d, d' \} / 2^n$$

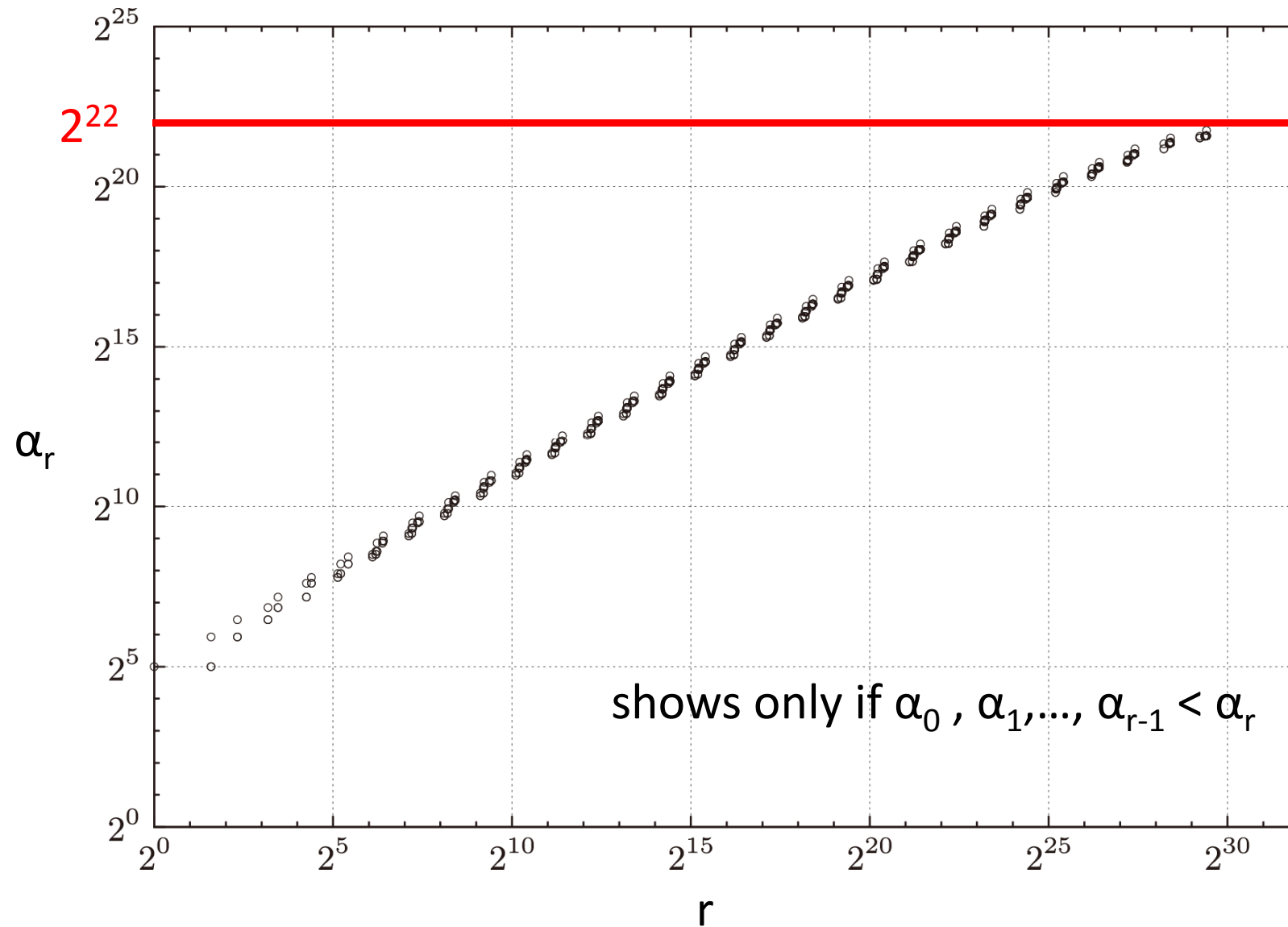
Approaches to Solve the Problem

- Make use of tools for the analysis of S-functions [Mouha, et al. '11, Leurent, '12]
- Our solution:
 - a recursive formula to compute α_r
 - Proposition
 - if $s_\ell = 0$, then $A_\ell = t_\ell A_{\ell-1} + B_{\ell-1}$
 - if $s_\ell \geq 1$, then $A_\ell = s_\ell B_\ell + A_{\ell-1}$
 - where $B_j = t_j A_{j-1} + B_{j-1}$ for $0 < j \leq \ell$, $A_j = s_j B_j + A_{j-1}$ for $0 < j \leq \ell-1$, $A_0 = 1$, and $B_0 = 0$
 - can be used to efficiently compute α_r

Graph of α_r



Graph of α_r



$$\alpha_{\max} = 3524578 \leq 2^{22}$$

- $\alpha_{\max} = 3524578$ is achieved when
 $r = 0x2aaaaaab, 0x55555555, 0xaaaaaaaaab, 0xd5555555$
- New version of [Lemma 3, MV04]

For any $0 \leq r \leq 2^{32}-1$,

$$\Pr_L[\text{Coll}_L(r, N, N')] \leq 2^{22} \max\{ d, d' \} / 2^n$$

where $d = \deg(\text{GHASH}_L(\varepsilon, N))$, $d' = \deg(\text{GHASH}_L(\varepsilon, N'))$

New Privacy Theorem

- $\text{Adv}_{\text{GCM}[\text{Perm}(n), \tau]}^{\text{priv}}(\mathcal{A}) \leq \frac{0.5(\sigma + q + 1)^2}{2^n} + \frac{2^{22}q(\sigma + q)(\ell_N + 1)}{2^n}$
- essentially the same as the original privacy bound
 - chosen plaintext attacks
 - main difference is 2^{22}
- If the nonce length is restricted to 96 bits, then

$$\text{Adv}_{\text{GCM}[\text{Perm}(n), \tau]}^{\text{priv}}(\mathcal{A}) \leq \frac{0.5(\sigma + q + 1)^2}{2^n}$$

New Authenticity Theorem

- $\text{Adv}_{\text{GCM}[\text{Perm}(n), \tau]}^{\text{auth}}(\mathcal{A}) \leq \frac{0.5(\sigma + q + q' + 1)^2}{2^n} + \frac{2^{22}(q + q' + 1)(\sigma + q)(\ell_N + 1)}{2^n} + \frac{q'(\ell_A + 1)}{2^\tau}$
- essentially the same as the original authenticity bound
 - main difference is 2^{22}
- If the nonce length is restricted to 96 bits, then

$$\text{Adv}_{\text{GCM}[\text{Perm}(n), \tau]}^{\text{auth}}(\mathcal{A}) \leq \frac{0.5(\sigma + q + q' + 1)^2}{2^n} + \frac{q'(\ell_A + 1)}{2^\tau}$$

Conclusions

- [Lemma 3, MV04] is not correct
 - the probability of a counter collision is higher than claimed
- The proofs can be repaired
 - new version of [Lemma 3, MV04]
 - new privacy theorem and new authenticity theorem
 - the bounds are worse than the original bounds, but GCM maintains the provable security (both in privacy and authenticity)
 - better bounds if the nonce length is restricted to 96 bits
- Open problem: Can we improve our security bounds?