Stronger Security Variants of GCM-SIV

Tetsu Iwata*[1] Kazuhiko Minematsu[2]

FSE 2017 Tokyo, Japan

March 8 2017

Nagoya University, Japan

NEC Corporation, Japan

* Supported in part by JSPS KAKENHI, Grant-in-Aid for Scientific Research (B), Grant Number 26280045.

Introduction

Nonce-Based AE and Its Limitation

- Nonce-based authenticated encryption : GCM [MV04], CCM [WHF02], OCB [RBBK01], EAX [BRW04], etc.
- They use a nonce for security: repeating the nonce has critical impact on security
  - Counter-then-MAC (incl. GCM): leaks plaintext difference
  - For GCM, even authentication key is leaked, allows universal forgery

---

[MV04] D.McGrew and J.Viega: The Security and Performance of the Galois/Counter Mode of Operation, Indocrypt 2004.

[WHF02] D.Whiting, R.Housley, and N.Ferguson: AES Encryption and Authentication Using CTR Mode and CBC-MAC. 2002.

[RBBK01] P.Rogaway, M.Bellare, J.Black, and T.Krovetz: OCB: A block-cipher mode of operation for efficient authenticated encryption. ACM CCS 2001.

[BRW04] M.Bellare, P.Rogaway, and D.Wagner: The EAX Mode of Operation. FSE 2004:

MRAE and SIV

Deterministic AE (DAE), a.k.a Misuse-resistant Nonce-based AE (MRAE) [RS06]

- Provides best-possible security if nonce is missing or exists but can be repeated by mistake
- Many concrete proposals including several CAESAR submissions

SIV, Synthetic IV [RS06]

- A general approach to construct MRAE
- use a PRF to generate IV (also used as a tag), use IV in IV-based encryption

---

[RS06] P.Rogaway and T.Shrimpton. A Provable-Security Treatment of the Key-Wrap Problem. Eurocrypt 2006.

How SIV works

Components:

- $\mathsf{F} : \mathcal{K} \times \mathcal{A} \times \mathcal{M} \to \mathcal{T}$
- $\mathsf{Enc} : \mathcal{K}' \times \mathcal{T} \times \mathcal{M} \to \mathcal{M}$, and the inverse, Dec
  - Typically a keystream generator

For encryption of plaintext M with associated data A:

1. $T \leftarrow \mathsf{F}_K(A, M)$
2. $C \leftarrow \mathsf{Enc}_{K'}(T, M)$
3. Return tag T and ciphertext C

Decryption: receives (A, T, C), computes $M \leftarrow \mathsf{Dec}_{K'}(T, C)$ and checks if $\mathsf{F}_K(A, M)$ matches with T

Provable security of SIV

We need PRF security of F and IV-based encryption security of Enc

GCM-SIV

GCM-SIV

- Proposed by Gueron and Lindell [GL15]
- Instantation of SIV using GCM components, GHASH and GCTR
  - Very fast AESNI implementations [GL15]
- Provable security $O(2^{(n-k)/2})$
  - Typically $n = 128$, $k = 32$. Thus about 48-bit security

Concrete Bound

For three-key version, with q encryption and q' decryption queries:

$$\mathbf{Adv}^{\mathrm{mrae}}_{\mathrm{GCM\text{-}SIV}}(\mathcal{A}) \leq 2\mathbf{Adv}^{\mathrm{prf}}_{E}(\mathcal{A}') + \frac{q^2}{2^{95}} + \frac{q^2 + q'}{2^{128}}$$

[GL15] S.Gueron and Y.Lindell : GCM-SIV: Full Nonce Misuse-Resistant Authenticated Encryption at Under One Cycle per Byte. ACM CCS 2015

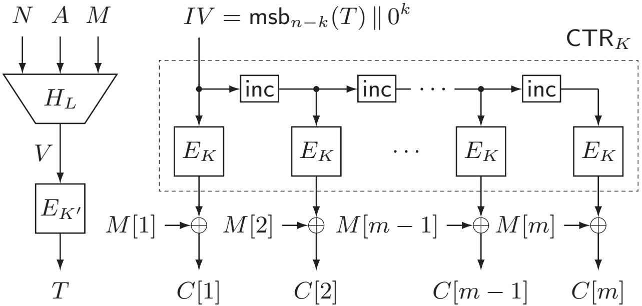
GCM-SIV

Specification:

Algorithm $\text{GCM-SIV-}\mathcal{E}_{\boldsymbol{K}}(N, A, M)$	**Algorithm** $\text{GCM-SIV-}\mathcal{D}_{\boldsymbol{K}}(N, A, C, T)$
1. $V \leftarrow H_L(N, A, M)$	1. $IV \leftarrow \mathsf{msb}_{n-k}(T) \parallel 0^k$
2. $T \leftarrow E_{K'}(V)$	2. $m \leftarrow \lvert C\rvert_n$
3. $IV \leftarrow \mathsf{msb}_{n-k}(T) \parallel 0^k$	3. $\mathbf{S} \leftarrow \mathsf{CTR}_K(IV, m)$
4. $m \leftarrow \lvert M\rvert_n$	4. $M \leftarrow C \oplus \mathsf{msb}_{\lvert C\rvert}(\mathbf{S})$
5. $\mathbf{S} \leftarrow \mathsf{CTR}_K(IV, m)$	5. $V \leftarrow H_L(N, A, M)$
6. $C \leftarrow M \oplus \mathsf{msb}_{\lvert M\rvert}(\mathbf{S})$	6. $T^* \leftarrow E_{K'}(V)$
7. **return** (C, T)	7. **if** $T \neq T^*$ **then return** $\perp$
	8. **return** M

- H_L is GHASH (with final xor of n-bit N)
  - $H_L(N, A, M) = \mathsf{GHASH}_L(A, M) \oplus N$
- CTR_K employs incrementation in the last k bits (as GCM)
  - Initial counter value is $\mathsf{msb}_{n-k}(T)$

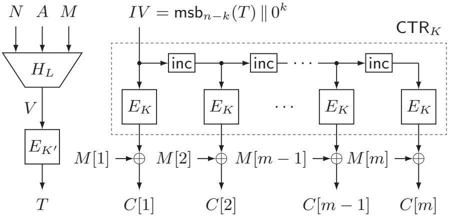
GCM-SIV



N A M
H_L
V
$E_{K'}$
T
$IV = \mathsf{msb}_{n-k}(T) \,\|\, 0^k$
CTR_K
inc
inc
inc
E_K
E_K
E_K
E_K
$M[1]$
$M[2]$
$M[m-1]$
$M[m]$
$C[1]$
$C[2]$
$C[m-1]$
$C[m]$

Security Bound is Tight

- Attack by counter collision search
- Fix A and M and make $2^{(n-k)/2}$ enc-queries (N_i, A, M) w/ distinct N_is
- For i and j w/ $\mathsf{msb}_{n-k}(T_i) = \mathsf{msb}_{n-k}(T_j)$, the adversary gets the same ciphertext



Considerations on Security

- Nonce-misuse-resistance : obivious quantitative gain in security from GCM
- While quantitatively the security can be degraded from GCM
  - distinguishing attack with $q = O(2^{(n-k)/2})$ queries
  - For GCM, there is no attack of the same complexity
    - if $|N| = 96$, IV is N itself – no counter collision
    - Even if $|N| \neq 96$ GCM bound is still good [NMI15]

[NMI15] : Y.Niwa, K.M., T.Iwata. GCM Security Bounds Reconsidered. FSE 2015.

Our Contributions

- The design strategy of reusing GCM components to build MRAE is practically valuable
- While the security offered by GCM-SIV may not be satisfactory in practice
- It seems some unexplored design space for stronger security
  - Up to the birthday bound ($n/2$-bit security)?
  - Beyond the birthday bound?

Our contributions

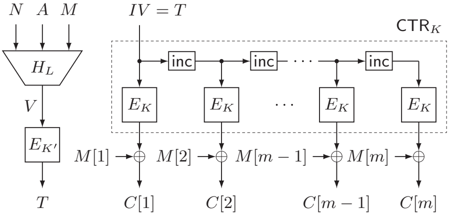
- GCM-SIV1: a minor variant of GCM-SIV achieving birthday bound security
- GCM-SIVr (for $r \geq 2$): by reusing r GCM-SIV1 instances to achieve $rn/(r+1)$-bit security

GCM-SIV1

GCM-SIV1

The changes are so simple:

- use the whole T as IV
- use full n-bit counter incrementation instead of k-bit incrementation



GCM-SIV1

Concrete Bound

If H_L is ϵ-almost universal (ϵ-AU),

$$\mathbf{Adv}^{\mathrm{mrae}}_{\mathrm{GCM\text{-}SIV1}}(\mathcal{A}) \leq 0.5q^2\epsilon + \frac{0.5q^2}{2^n} + \frac{\sigma^2}{2^n} + \frac{q}{2^n}$$

for q total (enc and dec) queries, each query is of length at most $n\ell$ bits, and σ queried blocks

If H_L is GHASH, $\epsilon = \ell/2^n$ thus $\ell q^2/2^n + \sigma^2/2^n + q/2^n$

Thus GCM-SIV1 is secure up to the standard birthday bound w.r.t. σ

Comparison of Bounds

Comprison of security bounds for GCM-SIV and GCM-SIV1

- Minimum attack complexity is increased ($(n-k)/2$ to $n/2$ bits)
- Still, depending on the average query length (σ/q), we can decribe two possible parameter settings where GCM-SIV1 beats GCM-SIV and vice versa

Implementation aspects

- GCM-SIV1 is very close to GCM-SIV, but
  - it needs full n-bit arithmetic addition
  - slightly degraded performance from GCM-SIV using GCTR

GCM-SIVr

Beyond the Birthday Bound (BBB)

Beyond $O(\sigma^2/2^n)$ bound – how ?

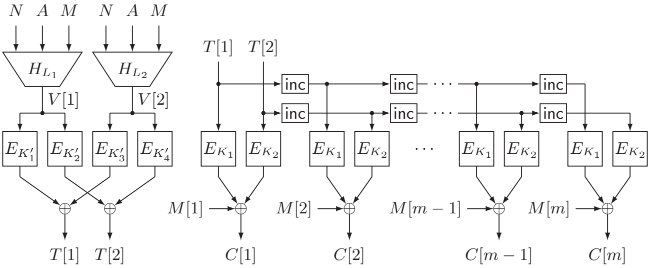
- Generic approach: use $2n$-bit blockcipher in SIV of $2n$-bit data path
- Effective instantiation not easy:
  - Widely-used 256-bit blockcipher?
  - Known constructions for $2n$-bit blockcipher from n-bit one (say, many-round Luby-Rackoff)
    - not fully efficient
    - not reusing GCM components (deviation from our strategy)

Our approach : GCM-SIVr

Compose r GCM-SIV1 instances in a manner close to black-box

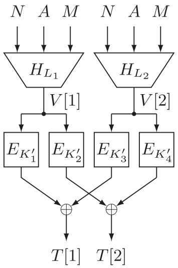
GCM-SIV2

1. Take two independently-keyed H_Ls to get $2n$-bit hash value $(V[1], V[2])$
2. Encrypt hash value with four blockcipher calls to get $2n$-bit tag $(T[1], T[2])$
3. Plaintext is encrypted by a sum of two CTR modes taking two IVs, $T[1]$ and $T[2]$



Proving Security of GCM-SIV2

- First game : Distinguish MAC function F2, which takes $(N, A, M) \to T$, from random function
  - Assuming blockciphers are random permutations



Analysis of F2

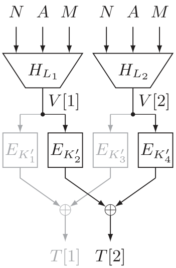
- SUM-ECBC by Yasuda [Y10] for BBB-secure PRF
- It is a sum of two Encrypted CBC-MACs (EMACs)
  - $T = E_{K_2}(\text{CBC-MAC}[E_{K_1}](M)) \oplus E_{K_4}(\text{CBC-MAC}[E_{K_3}](M))$
- [Y10] proved PRF bound $12\ell^4 q^3/2^{2n}$ for SUM-ECBC, thus $2n/3$-bit security (ignoring ℓ)

[Y10] K.Yasuda. The Sum of CBC MACs Is a Secure PRF. CT-RSA 2010

Analysis of F2

F2 is reduced to SUM-ECBC if

- output is chopped to n bits, either $T[1]$ or $T[2]$
- H_L is CBC-MAC
  - Osaki [O12] : CBC-MAC can be any ϵ-AU hash function



[O12] A.Osaki. A Study on Deterministic Symmetric Key Encryption and Authentication. Master's thesis, Nagoya University

Analysis of F2

Our task : extending [Y10][O12] so that F2 can handle $2n$-bit output

- Game-playing technique [BR06]
- [Y10][O12] employed a game having four cases
  - depending on the existance of collision in $V[i]$ for given input and for $i = 1, 2$
- We can employ a similar analysis as [Y10][O12] but need subcases to handle $2n$-bit output

PRF bound

$$\text{If } H_L \text{ is } \epsilon\text{-AU, } \mathbf{Adv}^{\mathrm{prf}}_{\mathsf{F2}}(\mathcal{A}) \leq \frac{8q^3}{3 \cdot 2^{2n}} + 6\epsilon^2 q^3$$

$$\text{If } H_L \text{ is GHASH, } \mathbf{Adv}^{\mathrm{prf}}_{\mathsf{F2}}(\mathcal{A}) \leq \frac{8.7\ell^2 q^3}{2^{2n}}$$

[BR06] M. Bellare, P. Rogaway: The Security of Triple Encryption and a Framework for Code-Based Game-Playing Proofs. EUROCRYPT 2006

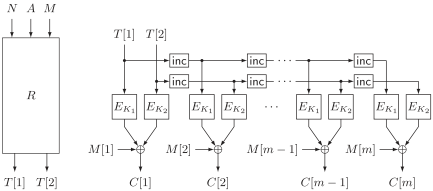
Analysis of Encryption Part

Second game: F2 is replaced with a random function R

- Encryption takes $2n$-bit random IV, $(T[1], T[2])$
- i-th counter block is $(T[1] + i - 1, T[2] + i - 1)$

Quite similar analysis as F2:

- $(N, A, M, i) \rightarrow (T[1] + i - 1, T[2] + i - 1)$ can be seen as a hashing process involving R and inc function
- Low collision probability for two distinct inputs, in fact $1/2^{2n}$



Security of GCM-SIV2

Concrete Bound of GCM-SIV2

For any (q, ℓ, σ)-adversary $\mathcal{A}$,

$$\mathbf{Adv}^{\mathrm{mrae}}_{\mathrm{GCM\text{-}SIV2}}(\mathcal{A}) \leq \frac{7\sigma^3}{2^{2n}} + 6\epsilon^2 q^3 + \frac{q}{2^{2n}},$$

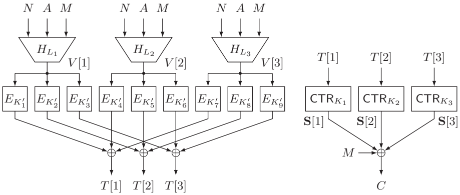
and if H_L is GHASH, the r.h.s. is bounded by

$$\frac{7\sigma^3}{2^{2n}} + \frac{6\ell^2 q^3}{2^{2n}} + \frac{q}{2^{2n}}.$$

Generalization to any r

The tag is generated by F$r : \mathcal{N} \times \mathcal{A} \times \mathcal{M} \to \{0,1\}^{nr}$.

- Analysis of Fr : we introduce $X = (x_1, \cdots, x_r) \in \{0,1\}^r$, where $x_i = 1$ indicates a collision on H_{L_i}'s outputs
- Exploit the symmetric property : the analysis is only depending on the Hamming weight of X
  - not much technical difficulty but needs careful work



Security of GCM-SIVr

- Let $f_{\mathsf{bad}}(p)$ be the probability of bad event invoked with weight of X being $p \in \{0, \ldots, r\}$
- Then $f_{\mathsf{bad}}(p)$ is bounded by $(2\epsilon)^r \cdot q^{r+1}$ for any $0 \leq p \leq r$

Concrete Bound of Fr

For any (q, ℓ, σ)-adversary $\mathcal{A}$,

$$\mathbf{Adv}^{\mathrm{prf}}_{\mathsf{F}r}(\mathcal{A}) \leq r \cdot 2^r \max_p \{f_{\mathsf{bad}}(p)\} \leq r \cdot (4\epsilon)^r \cdot q^{r+1},$$

which is $r \cdot (4\ell)^r \cdot q^{r+1}/2^{nr}$ if H_L is GHASH

Note: a dedicated analysis for given r can improve the bound constant (which we employed for $r = 2$) Encryption security is similarly derived as Fr

Concrete Bound of GCM-SIVr

For any (q, ℓ, σ)-adversary $\mathcal{A}$, we have

$$\mathbf{Adv}^{\mathrm{mrae}}_{\mathrm{GCM\text{-}SIV}r}(\mathcal{A}) \leq r \cdot (4\epsilon)^r \cdot q^{r+1} + \frac{4^r \cdot \sigma^{r+1}}{2^{nr}} + \frac{q}{2^{nr}},$$

and if GHASH is used for H_L,

$$\mathbf{Adv}^{\mathrm{mrae}}_{\mathrm{GCM\text{-}SIV}r}(\mathcal{A}) \leq \frac{r \cdot (4\ell)^r \cdot q^{r+1}}{2^{nr}} + \frac{4^r \cdot \sigma^{r+1}}{2^{nr}} + \frac{q}{2^{nr}}$$

Summary

GCM-SIVr is secure up to about $2^{rn/(r+1)}$ query complexity, and hence it asymptotically achieves full n-bit security

Conclusions

- Variants of GCM-SIV to offer quantitatively stronger security
- GCM-SIV1 : Standard $n/2$-bit security by tiny change to the original
- GCM-SIVr for $r \geq 2$: Use r GCM-SIV1 instances to go beyond the birthday bound, $rn/(r+1)$-bit security
  - Close to the black-box composition, highly parallel
  - (To our knowledge) the first concrete MRAE scheme to achieve asymptotically optimal security based on classical blockcipher
  - Large r implies large computation and large bandwidth, thus impractical

Conclusions

- Variants of GCM-SIV to offer quantitatively stronger security
- GCM-SIV1 : Standard $n/2$-bit security by tiny change to the original
- GCM-SIVr for $r \geq 2$: Use r GCM-SIV1 instances to go beyond the birthday bound, $rn/(r+1)$-bit security
  - Close to the black-box composition, highly parallel
  - (To our knowledge) the first concrete MRAE scheme to achieve asymptotically optimal security based on classical blockcipher
  - Large r implies large computation and large bandwidth, thus impractical

Thank you!