Stronger Security Variants of GCM-SIV

Tetsu Iwata\(^1\) \quad 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

---

References:


[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

How SIV works

Components:

- \( F : \mathcal{K} \times \mathcal{A} \times \mathcal{M} \to \mathcal{T} \)

- \( \text{Enc} : \mathcal{K}' \times \mathcal{T} \times \mathcal{M} \to \mathcal{M} \), and the inverse, \( \text{Dec} \)
  - Typically a keystream generator

For encryption of plaintext \( M \) with associated data \( A \):

1. \( T \leftarrow F_K(A, M) \)

2. \( C \leftarrow \text{Enc}_{K'}(T, M) \)

3. Return tag \( T \) and ciphertext \( C \)

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

Provable security of SIV
We need PRF security of \( F \) and IV-based encryption security of \( \text{Enc} \)
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:

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

GCM-SIV

Specification:

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCM-SIV-(\mathcal{E}_K(N, A, M))</td>
<td>GCM-SIV-(\mathcal{D}_K(N, A, C, T))</td>
</tr>
<tr>
<td>1. (V \leftarrow H_L(N, A, M))</td>
<td>1. (IV \leftarrow \text{msb}_{n-k}(T) \parallel 0^k)</td>
</tr>
<tr>
<td>2. (T \leftarrow E_K'(V))</td>
<td>2. (m \leftarrow</td>
</tr>
<tr>
<td>3. (IV \leftarrow \text{msb}_{n-k}(T) \parallel 0^k)</td>
<td>3. (S \leftarrow \text{CTR}_K(IV, m))</td>
</tr>
<tr>
<td>4. (m \leftarrow</td>
<td>M</td>
</tr>
<tr>
<td>5. (S \leftarrow \text{CTR}_K(IV, m))</td>
<td>5. (V \leftarrow H_L(N, A, M))</td>
</tr>
<tr>
<td>6. (C \leftarrow M \oplus \text{msb}_{</td>
<td>M</td>
</tr>
<tr>
<td>7. return ((C, T))</td>
<td>7. if (T \neq T^*) then return (\bot)</td>
</tr>
<tr>
<td>8. return (M)</td>
<td>8. return (M)</td>
</tr>
</tbody>
</table>

- \(H_L\) is GHASH (with final xor of \(n\)-bit \(N\))
  - \(H_L(N, A, M) = \text{GHASH}_L(A, M) \oplus N\)

- \(\text{CTR}_K\) employs incrementation in the last \(k\) bits (as GCM)
  - Initial counter value is \(\text{msb}_{n-k}(T)\)
GCM-SIV

\[ IV = \text{msb}_{n-k}(T) \parallel 0^k \]

\[ H_L \]

\[ \begin{array}{c}
N \quad A \quad M \\
\downarrow \quad \downarrow \quad \downarrow \\
H_L \\
\downarrow \\
V \\
\downarrow \\
E_K \\
\downarrow \\
T \\
\downarrow \\
E_K' \\
\end{array} \]

\[ \begin{array}{c}
M[1] \\
\oplus \\
C[1] \\
\downarrow \\
E_K \\
\downarrow \\
M[2] \\
\oplus \\
C[2] \\
\downarrow \\
\cdots \\
\downarrow \\
\cdots \\
\downarrow \\
E_K \\
\downarrow \\
M[m-1] \\
\oplus \\
C[m-1] \\
\downarrow \\
M[m] \\
\oplus \\
C[m] \\
\end{array} \]

\[ \text{CTR}_K \]
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 $N_i$s
- For $i$ and $j$ w/ $\text{msb}_{n-k}(T_i) = \text{msb}_{n-k}(T_j)$, the adversary gets the same ciphertext
Considerations on Security

- Nonce-misuse-resistance: obvious quantitative gain in security from GCM

- While quantitatively the security can be degraded from GCM
  - distinguishing attack with \( q = O\left(2^{(n-k)/2}\right) \) 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]

---

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-SIV$_r$ (for $r \geq 2$): by reusing $r$ GCM-SIV1 instances to achieve $rn/(r + 1)$-bit security
GCM-SIV1
The changes are so simple:

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

If $H_L$ is $\epsilon$-almost universal ($\epsilon$-AU),

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

for $q$ total (enc and dec) queries, each query is of length at most $n\ell$ bits, and $\sigma$ 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. $\sigma$
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 \((\sigma/q)\), we can describe two possible parameter settings where GCM-SIV1 beats GCM-SIV and vice versa
• 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-SIV

Beyond the Birthday Bound (BBB)

Beyond $O(\sigma^2/2^n)$ bound – how?

- Generic approach: use $2^n$-bit blockcipher in SIV of $2^n$-bit data path

- Effective instantiation not easy:
  - Widely-used 256-bit blockcipher?
  - Known constructions for $2^n$-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-SIV$_r$
Compose $r$ GCM-SIV1 instances in a manner close to black-box
1. Take two independently-keyed $H_L$s 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 $F_2$, which takes $(N, A, M) \rightarrow T$, from random function
  - Assuming blockciphers are random permutations
Analysis of $F_2$

- 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 $\ell$)

---

[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 $\epsilon$-AU hash function

---

Analysis of $F_2$

Our task: extending [Y10][O12] so that $F_2$ can handle $2^n$-bit output

- Game-playing technique [BR06]

- [Y10][O12] employed a game having four cases
  - depending on the existence 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 $2^n$-bit output

**PRF bound**

If $H_L$ is $\epsilon$-AU, $\text{Adv}_{F_2}^{\text{prf}}(A) \leq \frac{8q^3}{3 \cdot 2^{2n}} + 6\epsilon^2 q^3$

If $H_L$ is GHASH, $\text{Adv}_{F_2}^{\text{prf}}(A) \leq \frac{8.7\ell^2 q^3}{2^{2n}}$

Analysis of Encryption Part

Second game: $F_2$ 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 $F_2$:

- $(N, A, M, i) \rightarrow (T[1] + i - 1, T[2] + i - 1)$ 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 \((q, \ell, \sigma)\)-adversary \(A\),

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

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 $F_r : \mathcal{N} \times \mathcal{A} \times \mathcal{M} \rightarrow \{0, 1\}^{nr}$.

- Analysis of $F_r$: 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-SIV\_r

- Let $f_{\text{bad}}(p)$ be the probability of bad event invoked with weight of $X$ being $p \in \{0, \ldots, r\}$

- Then $f_{\text{bad}}(p)$ is bounded by $(2\epsilon)^r \cdot q^{r+1}$ for any $0 \leq p \leq r$

**Concrete Bound of F\_r**

For any $(q, \ell, \sigma)$-adversary $\mathcal{A}$,

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

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 $F_r$
Concrete Bound of GCM-SIV<sub>r</sub>
For any \((q, \ell, \sigma)\)-adversary \(A\), we have

\[
\text{Adv}_{\text{GCM-SIV}_r}^{\text{mrae}}(A) \leq r \cdot (4\epsilon)^r \cdot q^{r+1} + \frac{4^r \cdot \sigma^{r+1}}{2^{nr}} + \frac{q}{2^{nr}},
\]

and if GHASH is used for \(H_L\),

\[
\text{Adv}_{\text{GCM-SIV}_r}^{\text{mrae}}(A) \leq r \cdot (4\ell)^r \cdot q^{r+1} + \frac{4^r \cdot \sigma^{r+1}}{2^{nr}} + \frac{q}{2^{nr}}.
\]

Summary
GCM-SIV\(_r\) 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-SIV$_r$ 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-SIV$_r$ 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!