Finding GCM Weak Keys

Markku-Juhani O. Saarinen
mjos@reveresecurity.com

14 February 2011

FSE 2011 Rump Session

Why GCM is Relevant

NIST SP 800-8D, “Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC.” GCM is the approved authenticated encryption mode in NSA Suite B. Specifications exist for integration with the IPSec, TLS and SSH2 protocols.

Message Forgery

Let X be a concatenation of unencrypted authenticated data A, CTR-encrypted ciphertext C, and the lengths of A and C. GCM/GHASH uses Horner's rule to compute

$$Y_m = \bigoplus_{i=1}^{m} X_i \otimes H^i.$$

The final tag is $T = E_K(Y_m \oplus (IV \mid\mid 0^{31} \mid\mid 1))$.

If we know that $H^i = H^j$ with $i \neq j$, we may simply swap X_i and X_j and the resulting authentication tag stays the same.

Note that ciphertext is authenticated, not plaintext.

Let $o = ord(H)$ be the multiplicative order of H. Then $H^i = H^{i+o}$ for all i.

These observations are not accurate for GCM

D.A. MCGREW AND S. FLUHRER. “Multiple Forgery Attacks against Message Authentication Codes.”

McGrew and Fluhrer have observed in that once a single forgery has been performed, additional forgeries become easier; more specifically, the forgery probability for MAC algorithms such as CBC-MAC and HMAC increases cubically with the number of known text-MAC pairs, while for universal hash functions the forgery probability increases only **quadratically**.

H. HANDSCHUH AND B. PRENEEL. “Key-Recovery Attacks on Universal Hash Function based MAC Algorithms.”

Handschuh and Preneel have analyzed Key-Recovery Attacks on Universal Hash Function based MAC Algorithm. They give the number of weak keys in $GF(2^{128})$ as **one**. The design document of GCM only considers $H = 0$.

Cycle Length

Let g be a generator of $GF(2^{128})$ and i the index $g^i = H$. It is easy to see that $0 \leq i < 2^{128} - 1$ is essentially random for random K. If i divides the multiplicative group size $2^{128} - 1$, we get a shorter cycle.

The group order is quite smooth:

$$2^{128} - 1 = 3 \times 5 \times 17 \times 257 \times 641 \times 65537 \times 274177 \times 6700417 \times 67280421310721.$$

Hence there **are** large classes of weak keys K that produce cycles of length $o = 1, 3, 5, 15, 17$ etc.

Implication

Assume that K and therefore H are random and unknown to the attacker.

If we swap the X_0 and $X_{2^{32}-1}$ blocks then the forgery will be undetected with probability 2^{-96} rather than 2^{-128} as expected from a good MAC.

This is because $\gcd(2^{32}-1, 2^{128}-1) = 2^{32}-1$ and therefore $2^{-128+32} = 2^{-96}$ is the probability that H just happens to belong to this multiplicative subgroup.

Note that this does not violate the GCM security claim, which reduces a t-bit authentication forgery only to a $\sqrt{2^t}$ attack on the underlying block cipher!

Some very weak $H = E_K(0^{128})$ values

$o = 1$:

```
H = 80 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
```

$o = 3$:

```
H = 10 D0 4D 25 F9 35 56 E6 9F 58 CE 2F 8D 03 5A 94
H = 90 D0 4D 25 F9 35 56 E6 9F 58 CE 2F 8D 03 5A 94
```

$o = 5$:

```
H = 46 36 BD BD 1C 76 43 D3 4E E4 BB 1B F9 CA 08 4F
H = 92 17 8D 40 26 DA 1D CA 42 96 77 87 30 EB 9A 9E
H = 82 C7 C0 65 DF EF 4B 2C DD CE B9 A8 BD E8 C0 0A
H = D6 E6 F0 98 E5 43 15 35 D1 BC 75 34 74 C9 52 DB
```

Results: Finding bad keys in AES-128

TEST: Theorem. Iff the cycle o of H is divisible by d, then

$$H^{\frac{2^{128}-1}{d}} = 1.$$

This way we may find increasingly weak K values in AES-128:

$o \approx 2^{126.4150}$ $K =$ 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 02

$\cdots$

$o \approx 2^{96.0000}$ $K =$ 00 00 00 00 00 00 00 00 00 00 00 00 37 48 CF CE

$o \approx 2^{93.9352}$ $K =$ 00 00 00 00 00 00 00 00 00 00 00 00 42 87 3C C8

$o \approx 2^{93.4117}$ $K =$ 00 00 00 00 00 00 00 00 00 00 00 00 EC 69 7A A8

Is there a shortcut ?

Concluding..

It should be more widely recognized that there are classes of keys for which GCM/GHASH message authentication is **weak**. The "unit price" for GHASH collisions is low – similar feature to multicollision attacks. This should be taken into account when protocols are designed using these primitives.

It's apparent that $GF(2^{128}+12451)$ or $GF(2^{128}-15449)$ would be more secure fields than the cumbersome $GF(2^{128})$. These are Sophie Germain primes and hence the group order is not smooth.

Note that Bernstein's AES-Poly1305 uses $p = 2^{130} - 5$ and $p - 1 = 2 \times 23 \times 89706473951992278723018299378 3$, which is quite secure.

We are not aware of any method that maps weak H values to keys K in AES. Such methods may exist for other 128-bit block ciphers.