MGM Beyond the Birthday Bound

Denis Fomin and Alexey Kurochkin

June 5, 2019

- The Multilinear Galois Mode (MGM) is an authenticated encryption with associated data (AEAD) block cipher mode.
- It was originally proposed in[1] and was fully described later in[2].
- MGM mode was developed by the Technical Committee for standardization “Cryptography and Security Mechanism” (TC-26) and now is a prospective Russian standard of AEAD mode[3].

[1] Nozdrunov V., “Parallel and double block cipher mode of operation (PD-mode) for authenticated encryption”, CTCrypt 2017

[2] Nozdrunov V. and Shishkin V., “Multilinear Galois Mode (MGM)”, CFRG Draft, 2018, https://datatracker.ietf.org/doc/draft-smyshlyaev-mgm

[3] Federal Agency on Technical Regulating and Metrology, “Recommendations for standardization. Cryptography. Authentication encryption modes of block ciphers”, 2018, In Russian

Previous results

- In 2019 the MGM mode was analysed in the paradigm of provable security[4]
- That work shows that the privacy and authenticity of MGM mode is provably guaranteed (under security of the used block cipher) up to the birthday paradox bound.

[4] Akhmetzyanova L., Alekseev E., Karpunin G., and Nozdrunov V., "Security of Multilinear Galois Mode (MGM)" , Cryptology ePrint Archive, 2019/123 (2019), https://eprint.iacr.org/2019/123

Our goals

- At the same time no real attack has been published so far even in the unlimited amount of queries (even trivial ones).
- So our plan is:

Our goals

- At the same time no real attack has been published so far even in the unlimited amount of queries (even trivial ones).
- We want to propose two attacks that are based on birthday paradox.
- That means that these attacks do not threaten the security claims of the MGM mode.

Definition and Notations

Let V^n be the boolean (bit) vector space of dismention n.

For a vector $x \in V^n$ we call the value $|x| = n$ the length of the vector x.

We assume that any element of the vector space $x \in V_n$ can be represent as an element of the ring $\mathbf{Z}_{2^n}(+, \cdot)$ and as an element of a finite field $\mathbb{F}_{2^n}$ $(\oplus, \otimes)$.

Let $x, y \in V^{n/2}$ and $t \in \mathbf{Z}_{2^{n/2}}$:

$$(x\|y) \boxplus^l t = (x + t\|y)\,;\ (x\|y) \boxminus^l t = (x - t\|y)\,;$$
$$(x\|y) \boxplus^r t = (x\|y + t)\,;\ (x\|y) \boxminus^r t = (x\|y - t)\,.$$

Definition and Notations

- Let e be a block cipher with block length n and $K \in V^k$ be a key.
- Denote by $e_K(x)$ the encryption of a plaintext block x under the key K.
- The input of the MGM mode based on a cipher e is (K, N, P, A), where:
  - $K \in V^k$ – key ;
  - $N \in V^{n-1}$ – nonce ;
  - $P \in V^*$, $0 \leq |P| \leq 2^{n/2}$ – plain text ;
  - $A \in V^*$, $0 \leq |A| \leq 2^{n/2}$ – associated data.
- The output of the mode is $(N, A, C, , T)$, where:
  - $C \in V^*$, $|P| = |C|$ is a cipher text;
  - $T \in V^m$ is an authenticating tag.

Let's denote $\Big(|A| \| |C|\Big)$ as L and call it "length tag".

Encryption

Encryption

Encryption in MGM is a variant of CTR mode.

Initialization vector

$$Y_1 \ = \ e_K(0\|N)$$

Counter calculation

$$Y_i = Y_{i-1} \boxplus^r 1$$

Keystream calculation

$$G_i = e_K(Y_i)$$

Authenticating Tag Calculation

Initialization vector

$Z_1 \ = \ e_K(1\|N)$

Counter calculation

$Z_i = Z_{i-1} \boxplus^l 1$

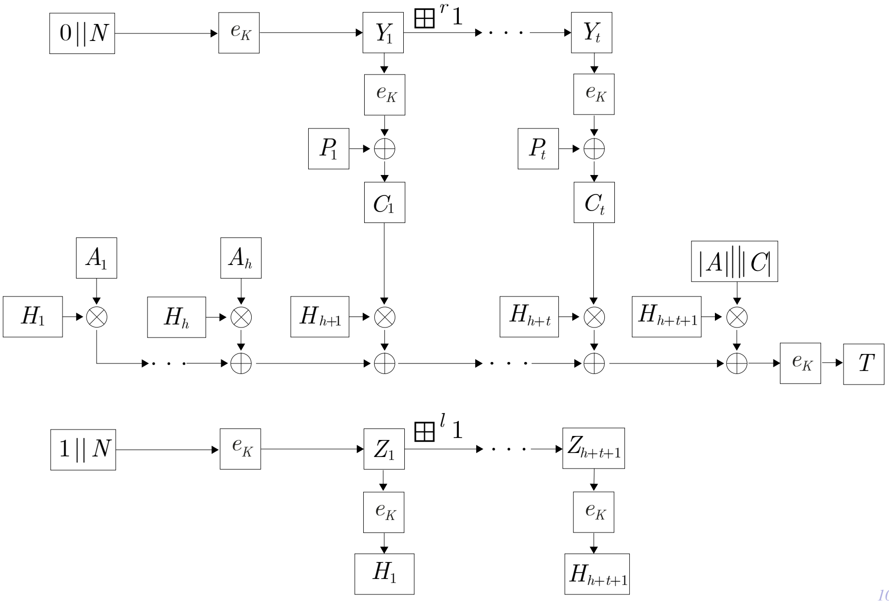
Additional value calculation

$H_i = e_k(Z_i)$

Authenticating Tag

$$T = e_K\left(\sum_{i=1}^{h} H_i \otimes A_i \oplus \sum_{j=1}^{q} H_{h+j} \otimes C_j \oplus H_{h+q+1} \otimes \Big(|A|\|\,|C|\Big)\right)$$

MGM mode


$0||N$
e_K
Y_1
$\boxplus^r 1$
$\cdots$
Y_t
e_K
e_K
P_1
P_t
C_1
C_t
A_1
A_h
$|A|||C|$
H_1
H_h
H_{h+1}
H_{h+t}
H_{h+t+1}
$\cdots$
$\cdots$
e_K
T
$1||N$
e_K
Z_1
$\boxplus^l 1$
$\cdots$
Z_{h+t+1}
e_K
e_K
H_1
H_{h+t+1}

Nonce reusing attack

If we have two different messages with the same authenticating tags and if we in addition have a possibility to authenticate an arbitrary message, it is possible to calculate the authenticating tag for the special message.

Nonce reusing attack

- Let we have two messages received using MGM mode under the same key K: (N_1, A_1, C_1, T_1), (N_2, A_2, C_2, T_2) and $T_1 = T_2$.
-

$$\sum_{i=1}^{h_1} H_{1,i} \otimes A_{1,i} \oplus \sum_{j=1}^{q_1} H_{1,h_1+j} \otimes C_{1,j} \oplus H_{1,h_1+q_1+1} \otimes L_1 =$$

$$= \sum_{i=1}^{h_2} H_{2,i} \otimes A_{2,i} \oplus \sum_{j=1}^{q_2} H_{2,h_2+j} \otimes C_{2,j} \oplus H_{2,h_2+q_2+1} \otimes L_2,$$

where L_1 and L_2 are the length tags and $L_1 = \left(n \cdot k_1^{(1)} \| n \cdot k_2^{(1)}\right)$, $L_2 = \left(n \cdot k_1^{(2)} \| n \cdot k_2^{(2)}\right)$.

Nonce reusing attack

$$\sum_{i=1}^{h_1} H_{1,i} \otimes A_{1,i} \oplus \sum_{j=1}^{q_1} H_{1,h_1+j} \otimes C_{1,j} \oplus H_{1,h_1+q_1+1} \otimes L_1 =$$

$$= \sum_{i=1}^{h_2} H_{2,i} \otimes A_{2,i} \oplus \sum_{j=1}^{q_2} H_{2,h_2+j} \otimes C_{2,j} \oplus H_{2,h_2+q_2+1} \otimes L_2,$$

If the left and the right sides of the equation above are multiplied by the same element α of the finite field $\mathbb{F}_2^n$ then we get the correct equation.

Nonce reusing attack

We can make the following message:

$$\begin{cases} A'_i = A_{1,i} \otimes \alpha, & 1 \leq i \leq h_1 \\ A'_{i+h_1} = C_{1,i} \otimes \alpha, & 1 \leq i \leq q_1 \end{cases}, \tag{1}$$

where α can be calculated from the equation:

$$L_1 \otimes \alpha = \left(n \cdot k_1^1 \| n \cdot k_2^1\right) \otimes \alpha = \left(0 \| n \cdot (k_1^1 + k_2^1)\right),\ k_i^j \in \mathbb{Z}, i, j \in \{1, 2\}.$$

Nonce reusing attack

We suppose that it's possible to request authenticating tag ("to sign") for the associated data $A' = A'_1 \| \ldots \| A'_{q_1+h_1}$:

$$T' = e_K\left(\left(\sum_{i=1}^{h_2} H_{2,i} \otimes A_{2,i} \oplus \sum_{j=1}^{q_2} H_{2,h_2+j} \otimes C_{2,j} \oplus H_{2,h_2+q_2+1} \otimes L_2\right) \otimes \alpha\right)$$

Nonce reusing attack

Let all the messages have the following structure: (N_i, A_i, C_i, T_i), where $|A_i| = |A_j|$, $|C_i| = |C_j|$, and all messages are calculated under the same key K.

1. Request D messages. With the probability $p \approx 1 - \exp\left\{-\frac{(D-1)^2}{2^{n+1}}\right\}$ two messages with numbers i and j such as $T_i = T_j$ will appear.
2. Make a new message from (N_i, A_i, C_i, T_i) using the equation (1).
3. Ask to authenticate this message.
4. Get the message $(K, N_i, \alpha \otimes (A_i \| C_i), T')$.
5. Make a new message with correct authenticated tag $(K, N_j, \alpha \otimes (A_j \| C_j), T')$.

How to get H_i?

Let's consider the following message (N_1, A_1, C_1, T_1), where $|A_1| = 0$, $C_1 = 0$, and $|C_1|$ is equal to 1. Then

$$T_1 = e_K(H_2 \otimes 1) = e_K\left(e_K\left(e_K\left(1||N_1\right) \boxplus^l 1\right)\right).$$

Let (K, N_2, A_2, C_2, T_2) be another message and $P_1 \oplus C_1 = e_K(Y_1) = e_K\left(e_K\left(0||N_2\right)\right)$ is equal to authenticating tag T_1. Then we can argue that:

$$e_K\left(e_K\left(e_K\left(1||N_1\right) \boxplus^l 1\right)\right) = e_K\left(e_K\left(0||N_2\right)\right) \Rightarrow e_K(1||N_1) = 0||N_2 \boxminus^l 1.$$

How to get H_i?

According to the MGM mode description $e_K\,(1\|N) = Z_1$.

Let $lsb_{\frac{n}{2}}(N_1) = lsb_{\frac{n}{2}}(N_2)$ there is such a value t: $t \in \mathbb{Z}$, $t < 2^{n/2}$:

$$Z_{t-1} = e_K(1||N_1) \boxplus^l t = 0||N_2 \boxplus^l (t-1) = (1||N_1),$$

and it is possible to calculate

$$H_{t-1} = e_K\,(Z_{t-1}) = e_K(e_K(1||N_1) \boxplus^l t) = e_K(1||N_1) = 0||N_2 \boxminus^l -1.$$

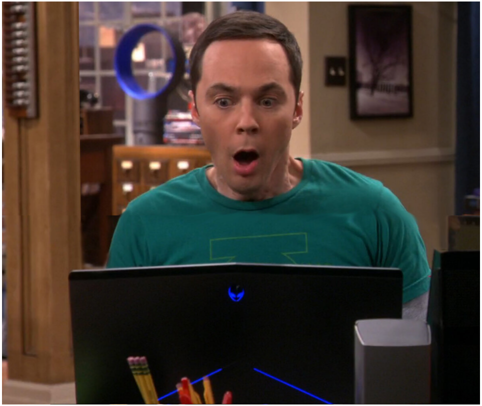
How to get H_i?

Let’s suppose that we have two different values $H_i = e_K(Z_i)$, $H_j = e_K(Z_j)$ and $e_K^{-1}(H_i)$, $e_K^{-1}(H_j)$ for some values $i < j < 2^{n/2}$.

Than means that

$$H_i = 0||N_i \boxminus^{l_i} -1, H_j = 0||N_j \boxminus^{l_j} -1.$$

And if $lsb_{\frac{n}{2}}(N_i) = lsb_{\frac{n}{2}}(N_j)$ then both H_i and H_j correspond to one IV.



Double H attack

- Let's suppose that we have two different values $H_i = e_K(Z_i), H_j = e_K(Z_j)$ and $e_K^{-1}(H_i)$, $e_K^{-1}(H_j)$ for some values $i < j < 2^{n/2}$.
- We also assume that $lsb_{\frac{n}{2}}(Z_i) = lsb_{\frac{n}{2}}(Z_j)$.
- Let $h, q \in \mathbb{N}_0$ and $h + q + 1 = j$ then we can form the following message S (value x will be determined later):

$$S = \left(\underbrace{0, 0, \ldots, 0}_{i-1}, x, \underbrace{0, 0, \ldots, 0}_{j-i-2}\right) = \left(\underbrace{A_1, \ldots, A_h, C_1, \ldots, C_q}_{j-1}\right).$$

Double H attack

- Let $h, q \in \mathbb{N}_0$ and $h+q+1 = j$ then we can form the following message S (value x will be determined later):

$$S = \left(\underbrace{0, 0, \ldots, 0}_{i-1}, x, \underbrace{0, 0, \ldots, 0}_{j-i-2}\right) = \left(\underbrace{A_1, \ldots, A_h, C_1, \ldots, C_q}_{j-1}\right).$$

- The authenticating tag T of the message S is calculated as follows:

$$T = e_K(x \otimes H_i \oplus L \otimes H_j),$$

where $L = (l(A) \| l(C))$ — length tag of message S.

- Fixing the values h and q we can calculate the value x using one of the following equations:

$$x \otimes H_i \oplus L \otimes H_j = e_K^{-1}(H_i);$$

$$x \otimes H_i \oplus L \otimes H_j = e_K^{-1}(H_j)$$

and authenticated tag will be equal to H_i and H_j respectively.

A pair of values h and q can be fixed by any of the j possible values and which means that we can calculate authenticating tag for $2 \cdot j$ messages without knowing the secret key K and moreover, half of these messages will have $T = H_i$ and the other half will have authenticated tag equal to H_j. That also means that in case of $j > 1$ we can also find a collision.

Double H attack

We suppose that all $lsb_{\frac{n}{2}}(N_i')$ and $lsb_{\frac{n}{2}}(N_i'')$ are equal.

1. Get m_1 messages (N_i', A_i', C_i', T_i'), where $|A_i'| = 0$, $|C_i'| = n$:

$$M_1 = \{Y_1(N_i)\}_{i=0}^{m_1} = \{e_K(e_K(0\|N_i))\}_{i=0}^{m_1}.$$

2. Get $2 \cdot m_2$ messages $(N_i'', A_i'', C_i'', T_i'')$, where $|A_i''| = 0$, $|C_i''| = 1$. We suppose that about the half of these messages is equal to zero $C_2'' = 0$ (one bit) and we have

$$M_2 = \{T_j\}_{j=0}^{m_2} = \left\{e_K\left(e_K\left(e_K\left(1\|N_i\right)\boxplus^{l}\right)\right)\right\}_{j=0}^{m_2}.$$

3. With some probability P_2 we find two equalities:

$$e_K(1\|N_1) = 0\|N_2 \boxminus^{l} 1, e_K(1\|N_3) = 0\|N_4 \boxminus^{l} 1,$$

$$lsb_{\frac{n}{2}}(N_1) = lsb_{\frac{n}{2}}(N_2) = lsb_{\frac{n}{2}}(N_3) = lsb_{\frac{n}{2}}(N_4).$$

We suppose that all $lsb_{\frac{n}{2}}\left(N_i'\right)$ and $lsb_{\frac{n}{2}}\left(N_i''\right)$ are equal.

4. Without loss of generality we suppose that $t_2 > t_1$. Fixing $h, q \in \mathbb{N}_0$ by any values such as: $h + q + 1 = t_2$ form the message:

$$S = (\underbrace{0, 0, \ldots, 0}_{i-1}, x, \underbrace{0, 0, \ldots, 0}_{t_2-t_1-2}) = \left(A_1, \ldots, A_h, C_1, \ldots, C_q\right),$$

   where x is calculated as follows:

$$x \otimes H_{t_1} \oplus L \otimes H_{t_2} = e_K^{-1}(H_{t_1}),\ L = (h \| q)\,.$$

5. The authenticating tag of message S is H_{t_1}.

To implement this attack we need:

- $m_1 + 2 \cdot m_2$ queries;
- memory $\mathcal{O}(m_1)$.

The probability that we can find at least one identical element in the sets M_1 and M_2 can be calculated as follows:

$$p_1 = 1 - \frac{\binom{2^n - m_1}{m_2}}{\binom{2^n}{m_2}} \approx 1 - \exp\left\{-\frac{m_1 m_2}{2^n}\right\}.$$

At the same time, exactly one identical element will be found with probability:

$$p = 2^n \frac{\binom{2^n-1}{m_1-1} \cdot \binom{2^n-m_1}{m_2-1}}{\binom{2^n}{m_1} \cdot \binom{2^n}{m_2}} \approx \frac{m_1 m_2}{2^n} \cdot \exp\left\{\frac{-1 + 2m_1 + 2m_2 - 2m_1 m_2}{2^{n+1}}\right\}.$$

And the required probability that more that one identical element will be found can be calculated using the equation $P_2 = p_1 - p$.

If there wasn't even one property, the attack would be inapplicable

We can implement the attack because:

- The first block of keystream is $e_k(e_K(0||N))$, but not $e_k(e_k(0||N_2) \boxplus^r 1))$.
- Next value Z_{t+1} is calculated as follows: $Z_{t+1} = e_K(Z_t \boxplus^l 1)$, but not $Z_{t+1} = e_K(Z_t \boxplus^r 1)$.
- Padding rule.
- Length tag is not encrypted.

Conclusion

In this paper we examined some aspects of the MGM AEAD mode and proposed two theoretical attacks that describe some properties of the studied mode.

Both attacks require about $\mathcal{O}\left(2^{n/2}\right)$ queries, with n the state size of used block cipher and do not threaten the security claims of MGM.

Authors would like to thank the reviewers for their helpful comments and efforts towards improving this work.